PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :		(11) International Publication Number:	WO 94/11987
H04N 1/46, 7/00, G09G 1/28	A1	(43) International Publication Date:	26 May 1994 (26.05.94)

(21) International Application Number: PCT/US93/11146

(22) International Filing Date: 16 November 1993 (16.11.93)

(30) Priority data:
103,763 16 November 1992 (16.11.92) IL

(71) Applicant: TECHNION RESEARCH & DEVELOPMENT FOUNDATION LTD. [IL/IL]; Senate House, Technion City, 32 000 Haifa (IL).

(71)(72) Applicant and Inventor: ZEEVI, Yehoshua, Y. [US/IL]; 48 Alexander Yanai Street, 34 816 Haifa (IL).

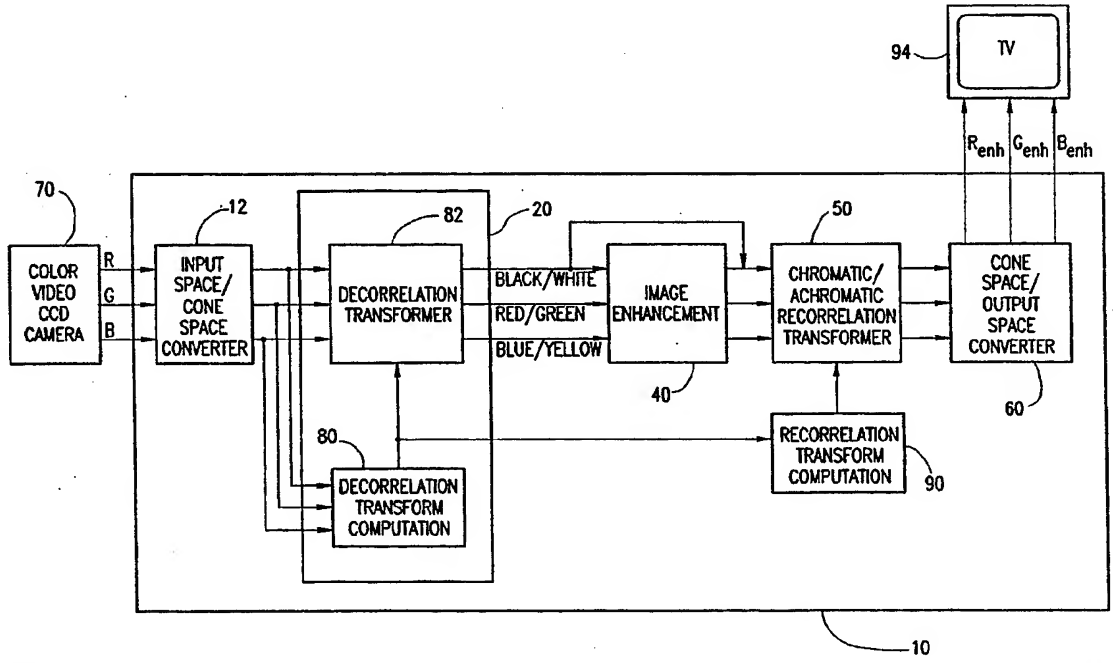
(72) Inventors: GINOSAR, Ran ; P.O. Box 104, 36 000 Nofit (IL). STUART, Wolf ; 12 Benyon Road, Pine Park, Johannesburg (ZA).

(74) Agents: DIPPERT, William, H. et al.; Cowan, Liebowitz & Latman, 605 Third Avenue, New York, NY 10158 (US).

(81) Designated States: AU, BB, BG, BR, CA, CZ, FI, HU, JP, KP, KR, LK, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SK, UA, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: APPARATUS AND METHOD FOR ENHANCING COLOR IMAGES



(57) Abstract

Color image enhancement apparatus (10) comprising apparatus (12) for receiving signals representing a color image, image processing apparatus, employing the received signals, for image processing of the high spatial frequency chromatic components of a color image, and apparatus for providing a color image from the output of said image processing apparatus.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

APPARATUS AND METHOD FOR ENHANCING COLOR IMAGES

FIELD OF THE INVENTION

The present invention relates to image processing generally and more particularly to detail enhancement in color images.

BACKGROUND OF THE INVENTION

Color image processing is discussed generally in the following texts, the disclosures of which are hereby incorporated by reference:

Rosenfeld, A. and Kak, A. C. Digital picture processing, Academic Press, 1982, and

Wysecki, G. and Stiles, W. S., Color science: concepts and methods, quantitiative data and formulae, John Wiley and Sons, 1982.

SUMMARY OF THE INVENTION

The present invention seeks to provide an improved system for enhancing color images which is based on an adaptive opponent color model.

The present invention also seeks to provide a system for enhancing color images which takes into account properties of the human visual system (HVS). Models of the human vision system are described in Chapters 1 and 2 of Appendix A, appended hereto.

The present invention also seeks to provide a system for adapting a color image for perception by color blind individuals.

There is thus provided, in accordance with a preferred embodiment of the present invention, color image enhancement apparatus including apparatus for receiving signals representing a color image, image processing apparatus, employing the received signals, for image processing of the high spatial frequency chromatic components of the color image, and apparatus for providing a color image using the output of the image processing apparatus.

Further in accordance with a preferred embodiment of the present invention, the apparatus for providing a color image also employs an achromatic component of the color image which was not processed by the image processing apparatus.

Still further in accordance with a preferred embodiment of the present invention, the apparatus for providing a color image also employs an achromatic component of the color image

which was processed by the image processing apparatus.

Still further in accordance with a preferred embodiment of the present invention, the apparatus for providing a color image also employs an achromatic component of the color image.

Additionally in accordance with a preferred embodiment of the present invention, the apparatus for receiving includes apparatus for receiving signals representing a color image in a first color space, and apparatus for transforming the received signals from the first color space into a color space which simulates the cones of the human visual system.

Further in accordance with a preferred embodiment of the present invention, the first color space includes an RGB space and the cone-simulating color space includes a (V_l, V_m, V_s) color space.

Additionally in accordance with a preferred embodiment of the present invention, the apparatus for providing includes apparatus for transforming the output of the image processing apparatus from the cone-simulating color space to a second color space.

Still further in accordance with a preferred embodiment of the present invention, the second color space includes an RGB space.

Additionally in accordance with a preferred embodiment of the present invention, the apparatus for receiving includes apparatus for at least partially decorrelating signals representing the color image.

Further in accordance with a preferred embodiment of the present invention, the apparatus for decorrelating includes

apparatus for performing a Karhunen-Loeve transform.

Still further in accordance with a preferred embodiment of the present invention, the apparatus for providing includes apparatus for performing an inverse of the Karhunen-Loeve transform.

There is also provided, in accordance with another preferred embodiment of the present invention, color image enhancement apparatus including apparatus for image processing of the high spatial frequency chromatic components of a color image.

There is further provided, in accordance with another preferred embodiment of the present invention, color image enhancement apparatus including apparatus for image processing of the high spatial frequency chromatic components of a color image, whereby a resulting enhanced color image is produced at least mainly from the high spatial frequency components.

Further in accordance with a preferred embodiment of the present invention, the high spatial frequency components are those contained in the top half, top third, or top quarter of the spatial frequency range of the color image.

Further in accordance with a preferred embodiment of the present invention, the power spectrum of the high spatial frequency chromatic components is similar to that of a high spatial frequency achromatic component of the color image.

There is also provided, in accordance with yet a further preferred embodiment of the present invention, color image enhancement apparatus for modifying a color image for perception by a color-blind individual, the apparatus including

apparatus for receiving signals representing the color image, image processing apparatus including apparatus, employing the received signals, for modifying the color image such that at least one color in the color image which a color-blind individual cannot differentiate is transformed to at least one color which the color-blind individual can differentiate, and apparatus for providing a color image using the output of the image processing apparatus.

Further in accordance with a preferred embodiment of the present invention, apparatus for identifying at least one color which the color-blind individual does not differentiate is also provided.

Further in accordance with a preferred embodiment of the present invention, the apparatus for modifying includes apparatus, employing the received signals, for modifying the color image such that a color in the color image which a dichromate does not differentiate is transformed to at least one color which the dichromate can differentiate.

Still further in accordance with a preferred embodiment of the present invention, the apparatus for modifying includes apparatus, employing the received signals, for modifying the color image such that at least one color in the color image which a monochromate does not differentiate is transformed to a color which the monochromate can differentiate.

Additionally in accordance with a preferred embodiment of the present invention, the image processing apparatus includes apparatus for enhancing chromatic differences between spatially adjacent colors.

Further in accordance with a preferred embodiment of the present invention, the apparatus for enhancing chromatic differences includes apparatus for enhancing chromatic differences between spatially adjacent colors which are achromatically indifferentiable.

Still further in accordance with a preferred embodiment of the present invention, the apparatus for enhancing chromatic differences includes apparatus for enhancing chromatic differences between spatially adjacent colors which are achromatically differentiable.

Additionally in accordance with a preferred embodiment of the present invention, apparatus is also provided for enhancing primarily a high spatial frequency portion of an achromatic component of the color image.

Further in accordance with a preferred embodiment of the present invention, apparatus is also provided for enhancing in the presence of noise including apparatus for enhancing primarily a low spatial frequency portion of an achromatic component of the color image.

There is also provided, in accordance with another preferred embodiment of the present invention, a method for color image enhancement including the steps of receiving signals representing a color image, employing the received signals for image processing of the high spatial frequency chromatic components of the color image, and providing a color image using the output of the apparatus for image processing.

Further in accordance with a preferred embodiment of

the present invention, the step of providing a color image also employs an achromatic component of the color image which was not processed in the image processing step.

Still further in accordance with a preferred embodiment of the present invention, the step of providing a color image also employs an achromatic component of the color image which was processed in the image processing step.

Additionally in accordance with a preferred embodiment of the present invention, the step of providing a color image also employs an achromatic component of the color image.

Further in accordance with a preferred embodiment of the present invention, the step of receiving includes the steps of receiving signals representing a color image in a first color space, and transforming the received signals from the first color space into a color space which simulates the cones of the human visual system.

Still further in accordance with a preferred embodiment of the present invention, the first color space includes an RGB space and the cone-simulating color space includes a (V_l, V_m, V_s) color space.

Further in accordance with a preferred embodiment of the present invention, the step of providing includes the step of transforming the output of the employing step from the cone-simulating color space to a second color space.

Still further in accordance with a preferred embodiment of the present invention, the second color space includes an RGB space.

Additionally in accordance with a preferred embodiment

of the present invention, the step of receiving includes the step of at least partially decorrelating signals representing the color image.

Further in accordance with a preferred embodiment of the present invention, the step of decorrelating includes the step of performing a Karhunen-Loeve transform.

Still further in accordance with a preferred embodiment of the present invention, the step of providing includes the step of performing an inverse of the Karhunen-Loeve transform.

There is also provided, in accordance with another preferred embodiment of the present invention, a color image enhancement method including the step of image processing of the high spatial frequency chromatic components of a color image.

There is further provided, in accordance with another preferred embodiment of the present invention, a color image enhancement method including the step of image processing of the high spatial frequency chromatic components of a color image, whereby a resulting enhanced color image is produced at least mainly from the high spatial frequency components.

Further in accordance with a preferred embodiment of the present invention, the high spatial frequency components include the top half, top third or top quarter of the spatial frequency range of the color image.

Still further in accordance with a preferred embodiment of the present invention, the power spectrum of the high spatial frequency chromatic components is similar to that of a high spatial frequency achromatic component of the color image.

Additionally in accordance with a preferred embodiment of the present invention, there is provided a color image enhancement method for modifying a color image for perception by a color-blind individual, the method including the steps of receiving signals representing the color image, image processing the color image, including the step of employing the received signals for modifying the color image such that at least one color in the color image which a color-blind individual cannot differentiate is transformed to at least one color which a color-blind individual can differentiate, and providing a color image using the output of the apparatus for image processing.

Further in accordance with a preferred embodiment of the present invention, the method also includes the step of identifying at least one color which the color-blind individual does not differentiate.

Still further in accordance with a preferred embodiment of the present invention, the step of employing includes the step of employing the received signals for modifying the color image such that a color in the color image which a dichromate does not differentiate is transformed to at least one color which the dichromate can differentiate.

Additionally in accordance with a preferred embodiment of the present invention, the step of employing includes the step of employing the received signals for modifying the color image such that at least one color in the color image which a monochromate does not differentiate is transformed to a color which the monochromate can differentiate.

Further in accordance with a preferred embodiment of

the present invention, the step of image processing includes the step of enhancing chromatic differences between spatially adjacent colors.

Still further in accordance with a preferred embodiment of the present invention, the step of enhancing chromatic differences includes the step of enhancing chromatic differences between spatially adjacent colors which are achromatically indifferentiable.

Still further in accordance with a preferred embodiment of the present invention, the step of enhancing chromatic differences includes the step of enhancing chromatic differences between spatially adjacent colors which are achromatically differentiable.

Additionally in accordance with a preferred embodiment of the present invention, the method includes the step of enhancing primarily a high spatial frequency portion of an achromatic component of the color image.

Still further in accordance with a preferred embodiment of the present invention, the method includes a step of enhancing in the presence of noise including the step of enhancing primarily a low spatial frequency portion of an achromatic component of the color image.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will be understood and appreciated from the following detailed description, taken in conjunction with the drawings in which:

Fig. 1 is a simplified block diagram of image enhancement apparatus constructed and operative in accordance with a first preferred embodiment of the present invention; and

Fig. 2 is a simplified block diagram of image enhancement apparatus constructed and operative in accordance with a second preferred embodiment of the present invention.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Reference is now made to Fig. 1 which illustrates image enhancement apparatus 10 constructed and operative in accordance with a first preferred embodiment of the present invention.

Image enhancement apparatus 10 comprises an input space/cone space converter 12, a decorrelator 20, an image enhancement unit 40, a chromatic/achromatic recorrelator 50 and a cone space/output space converter 60.

The image enhancement apparatus 10 receives signals representing a color image in a color space such as RGB or YC_rC_b. The source of these signals may, for example, be a color video CCD camera 70.

Preferably, the input signals are converted by input space/cone space converter 12 to a representation similar to that employed by HVS cones, which are biological retinal receptors. One example of a cone-simulating color space is a (V_l, V_m, V_s) space, as described in Appendix A. However, any other suitable cone-simulating color space may be employed.

The output of input space/cone space converter 12 is provided to decorrelator 20 which is operative to reduce the correlations between the plurality of components of the output of converter 12.

According to a preferred embodiment of the present invention, decorrelator 20 is based upon the Karhunen-Loeve transform (KLT), which is described in the following publication, the disclosure of which is incorporated herein by reference:

Gonzalez, R. C. and Wintz, P. Digital image

processing, Addison-Wesley, 1987, pp. 122-130.

Decorrelator 20 preferably comprises a decorrelation transform computation unit 80 and a decorrelation transformer 82. Decorrelation transform computation unit 80 computes the transformation to be applied to the output of converter 12, as described in more detail on page 8 of the document appended hereto and referenced Appendix B. Decorrelation transformer 82 applies the transformation to the converter 12 output.

Alternatively, decorrelation transform computation unit 80 may be eliminated. Instead, a KLT transform may be precomputed based on a sample of images which is preferably representative of the type of images for which it is desired to employ the apparatus of Fig. 1. For example, as explained in more detail in Appendix B, a KLT transform suitable for daylight landscapes may be computed based upon a suitable sample including a variety of daylight landscapes.

In the illustrated embodiment, the components of the output of decorrelator 20, preferably comprising cone-simulating, decorrelated image data, are labelled black/white, red/green and blue/yellow, respectively, to facilitate an intuitive and rapid understanding of the present invention. However, it is appreciated that the output channels of decorrelator 20 may not, in fact, correspond exactly to this verbal description.

The cone-simulating, decorrelated image data, also termed herein (K_1, K_2, K_3), is provided to image enhancement unit 40. A particular feature of the present invention is that image enhancement unit 40 preferably processes high spatial frequency chromatic components of the color image data and that

the color image data output by image enhancement unit 40 is preferably produced mainly or entirely by processing of the high spatial frequency chromatic components.

In the present specification, the term "high spatial frequency chromatic components" refers to chromatic components whose spatial frequency is included in the top half, third or quarter, on a log scale, of the spatial frequency range of the entire color image. For example, if the size of the color image is 512 pixels x 512 pixels, the maximum spatial frequency along a particular dimension of the color image is 512/2 = 256 cycles per image, and any of the following categories of chromatic components may be regarded as "high spatial frequency chromatic components":

"top half" chromatic components whose spatial frequency is greater than $256^{1/2}$ = 16 cycles per image;

"top third" chromatic components whose spatial frequency is greater than $256^{2/3}$ = approximately 40 cycles per image; and

"top quarter" chromatic components whose spatial frequency is greater than $256^{3/4}$ = 64 cycles per image.

Alternatively, the term "high spatial frequency chromatic components" refers to those chromatic components having a power spectrum resembling the power spectrum of high spatial frequency achromatic components of the color image. The term "high spatial frequency achromatic components" may be defined analogously to the previous definitions in which "high spatial frequency chromatic components" were defined independent of

achromatic components.

An operational definition of the term "high spatial frequency chromatic components" in applications in which high-pass achromatic enhancement is provided, is those chromatic components which survive filtering with the same high-pass filters employed for the achromatic enhancement.

Preferably, color enhancement unit 40 is operative to enhance details in the color image by enhancing chromatic differences between spatially adjacent colors which may or may not be achromatically differentiable.

Preferably, the color enhancement unit 40 is also operative to process achromatic components of the color image. According to one alternative embodiment of the present invention, primarily the high frequency portion of the achromatic components is processed.

Alternatively, the achromatic processing performed by color enhancement unit 40 primarily comprises processing of the low frequency portion of the achromatic components of the color image. This embodiment is particularly useful in the presence of strong noise, because empirical findings indicate that most of the noise occurs in the achromatic components.

Alternatively, in the presence of noise, even the low pass filtering of the achromatic components may be eliminated and processing may be restricted only to the chromatic components of the color image.

The output of color image enhancement unit 40 is provided to recorrelating unit 50 which performs the inverse of the transformation applied by decorrelating unit 82. The inverse

transform is computed by inverse transform computation unit 90 which receives an indication of the original transform from decorrelating transform computation unit 80.

Alternatively, if decorrelating transform computation unit 80 is eliminated, as explained above, recorrelation transform computation unit 90 may be eliminated and the inverse transform employed by recorrelating unit 50 may be predetermined, using a representative set of color images, as explained above.

The output of recorrelating unit 50 is provided to cone space/output space converter 60 which is operative to convert the data from cone space representation to output space representation. Any suitable output space may be employed, such as the RGB space of a CRT. The output space may or may not be the same as the input space in which the data received by chromatic enhancement apparatus 10 is represented.

The output of the apparatus 10 is an enhanced color image, on any suitable medium such as a CRT 94.

A detailed description of computational features of the various components of the apparatus of Fig. 1 appears in Appendix B. Other features of the apparatus of Fig. 1 are described in Appendix A.

It is appreciated that, in actual practice, many modifications are possible. For example, of course, instead of providing a sequence of units for performing a sequence of transformations on input color image data, a single unit may be provided which performs a single transformation which is the composition of the sequence of transformations. In other words,

the result of applying the single transformation is the same as the result of sequentially applying the sequence of transformations.

Reference is now made to Fig. 2 which illustrates a system for adapting color images for perception by color blind individuals, and which is based on the apparatus described above with reference to Fig. 1. Substantially identical elements in Figs. 1 and 2 are identically numbered for easy reference and are not described again hereinbelow for brevity.

In Fig. 2, a color image processing unit 100 is operative to "cross-enhance" at least one chromatic channel of the color image which is not perceived by a color blind individual using at least one other channel of the color image which is perceived by the color blind individual. For example, in the case of an individual who does not differentiate red/green, edges in the red/green channel will be enhanced in either or both of the blue/yellow and black/white channels.

A preferred method of operation for color image processing unit 100 is as follows:

a. Diagnose the type of color blindness of a target individual, preferably employing a computerized color blindness diagnosis system 110 interacting with the color blind individual. The computerized color blindness diagnosis system 110 is operative to determine which chromatic channels are not perceived and may be based on any suitable conventional testing materials for diagnosing color blindness. Color blindness diagnosis system 110 may, for example, be based upon the the following publication and the publications cited thereby, the

disclosures of which are hereby incorporated by reference:

Rushton, W. A. H., "Visual pigments and color blindness", Scientific American, Vol. 232(3), March 1975, pp. 64 - 74.

Alternatively, the color blindess diagnosis information may be provided manually or from a memory in which color blindness diagnoses were prestored.

According to still a further alternative, the operation of image processing unit 100 may be customized to adapt color images for perception only by a predetermined type of color blind individual, such as a red/green color blind individual.

In the illustrated embodiment, the image processing unit 100 is illustrated as "cross-enhancing" the red/green channel by "transferring" red/green information to other channels, namely the blue/yellow channel and/or the achromatic channel. It is appreciated, however, that image processing unit 100 may be operative to "cross-enhance" any one or any subset of the channels received thereby.

b. A gradient computation unit 120 is employed to detect gradients in the red/green channel.

c. The detected gradients are transferred to the blue/yellow and/or achromatic channels by a gradient application unit 130. For example, the gradient or spatial derivative detected for the red/green channel is preferably multiplied by a calibrating constant and may then be added to either or both of the blue/yellow and black/white channels.

It will be appreciated by persons skilled in the art

that the present invention is not limited to what has been particularly shown and described hereinabove. Rather, the scope of the present invention is defined only by the claims that follow:

Contents

List of Figures

List of Tables

Abstract

We investigate an opponent colour model of the HVS (human visual system) and show how this model can be used for colour image enhancement, in particular, colour edge enhancement.

The principle of the model is that in the HVS cone tristimuli are transformed to three opponent channels (black-white, red-green and blue-yellow) so that redundant information is minimized and energy compressed. Therefore after transforming a digital colour image represented by standard R, G and B tristimuli to cone tristimuli we use a Karhunen-Loéve transform to simulate the HVS opponent transform.

The model produces one all positive channel that contains at least 98 % of the total image energy. The spectral sensitivity function of this channel is very similar to the CIE luminous efficiency channel and thus contains the achromatic information. The model also results in two other antagonistic channels whose spectral sensitivity functions are similar to the Red-Green and Blue-Yellow opponent chromatic channels present in the HVS. The achromatic channel contains all spatial frequencies from dc upwards while the two chromatic channels are essentially lowpass.

For a set of similar images the model may be simplified from the general case of a Karhunen-Loéve transform per image to a simple linear matrix transformation. This is shown for a set as general as the set of natural daylight images.

Based on recent evidence of opponent centre-surround red-green and blue-yellow retinal receptive fields we enhance chromatic information by creating colour Mach bands. This is done by convolving the chromatic channels with a digital Laplacian mask. The resultant zero crossings are then scaled by a parameter, α, and added to the original channel.

The result is a sharper image with small colour detail highlighted and colours at least as rich as in the original image. Additional detail that achromatic edge enhancement does not affect is improved. This result is contrary to popular belief that since the chromatic components contain so little information, they need not be enhanced and can in fact be blurred. Because of limitations of the HVS very high frequency enhancement cannot be obtained using chromatic edge enhancement alone and the best result is obtained when combining standard achromatic enhancement with the novel chromatic edge enhancement. Recommended values of the parameter α are $\alpha = 1$ in the achromatic channel and $\alpha = 3$ in the chromatic channels.

Because of the spatial frequency properties of the model channels, most high frequency

image noise is in the achromatic channel. Edge enhancement on the chromatic channels does not perceptibly increase the noise and thus can be used without fear of degrading the image.

The results show that chromatic edge enhancement provides enhancement not possible using achromatic edge enhancement alone and that it may be used without degrading the image even in the presence of high frequency noise. In addition, the simplified model and enhancement procedure proposed are both suitable for real time imaging systems.

List of Symbols and Abbreviations

α_i — Enhancement parameter of channel i
C_x — Autocovariance matrix of ensemble X
CCD — Charge coupled device
CSF — Contrast Sensitivity Function
$d(i,j)$ — 3×3 unit impulse centred at (i,j)
e_{ij} — j'th element of i'th normalized eigenvector of C_x
Γ — Heuristic measure of post processing image improvement
Φ — Opponent transform matrix
Φ_c — Constant Opponent transform Matrix of ensemble
HVS — Human Visual System
HSI — Hue, Intensity and Saturation space
$i(\lambda)$ — incident radiant energy
$I(\lambda)$ — Radiant energy perceived at the retina
K_1 — Black-White (Achromatic) channel
K_2 — Red-Green opponent chromatic channel
K_3 — Blue-Yellow opponent chromatic channel
K_{in} — Noisy version of K_i channel
KLT — Karhunen-Loéve Transform
λ_i — i'th eigenvalue of C_x
L_n — Number of rows of image n
L_8 — 3×3 digital Laplacian
LGN — Lateral Geniculate Cortex
LMS — Long, Medium and Short wavelength cone responses
LPC — Linear Predictive Coding
m_x — mean vector of ensemble x
M — number of spectral planes in multispectral image
N_i — average noise power in channel i
NA — 3×3 neighborhood average mask
P_i — Average power in i'th spectral plane
PSD — Power Spectral Density

ρ — fraction of occupied colour space
RGB — Red, Green and Blue tristimuli of digital colour image
$S_L(\lambda), S_M(\lambda), S_S(\lambda)$ — Colour matching functions yielding LMS tristimuli
$S(k,l)$ — Saturation at point (k,l) in image
$\bar{S}_3$ — Average saturation in 3×3 neighborhood
T_{RV} — Linear transformation from RGB space to Vos-Walraven cone space
T_{RX} — Linear transformation from RGB space to 1931 CIE XYZ space
$T_{V\bar{X}}$ — Linear transformation from Vos-Walraven cone space to CIE $\bar{X}\bar{Y}\bar{Z}$ space
$V_l(\lambda), V_m(\lambda), V_s(\lambda)$ — Vos-Walraven cone spectral responses
W_n — Number of columns of image n
$\bar{W}$ — Average Weber fraction for entire image
$\bar{W}(k,l)$ — Estimated Weber fraction at point (k,l)
XYZ — CIE 1931 imaginary tristimulus system
$\bar{X}\bar{Y}\bar{Z}$ — CIE 1964 modified imaginary tristimulus system
$\mathbf{X}$ — Matrix of intensity vectors $X(k,l)$
$X(k,l)$ — $1 \times M$ intensity vector at location (k,l) across the M spectral planes
x_i — intensity at (k,l) for plane i

Chapter 1

Introduction

The research presented in this report is concerned with the enhancement of digital colour images and as such is also concerned with investigating the appropriate enhancement environment.

When speaking of colour image enhancement it is necessary to make the distinction between enhancement for image understanding purposes, for example satellite images, and for visual quality improvement such as television images. In the first instance actual colours are insignificant and may be altered at will to make certain features more prominent. In the second type of enhancement, with which we are concerned, true colours must be maintained.

There are two main ways in which a colour image can be visually enhanced. The first way is to alter the image such that undesirable colour casts are removed and the scene is rendered as if illuminated by a neutral white illuminant. In the literature this is referred to as the problem of colour constancy. The other main branch of colour image enhancement is enhancement of detail, noise removal, deblurring as well as other well known image enhancement techniques.

Although we are not concerned with colour constancy in this study we will discuss it briefly as it does introduce some of the techniques of dealing with colour images. The problem of colour constancy is related to how the HVS (human visual system) is consistently able to see the same colours in a scene even under different conditions of illumination whereas a camera used under similar conditions would result in an image with distorted colours.

To understand colour constancy it has been assumed that the HVS somehow calculates the reflectance, an intrinsic property of a material, at each point in the scene and in this way disregards the illuminant. In Buchsbaums spatial processor model [1] a finite dimensional linear model of illuminance and reflectance is used to estimate first the illuminance at each point and from this the reflectance. Gershon and Jepson extended Buchsbaums model such that the illuminance and reflectance basis functions are based on statistical measurements of naturally occurring reflectances and illuminances [2]. A similar linear model approach has also been proposed by Maloney and Wandell [3].

Another approach to this problem is given by Lands Retinex theory [4–6]. Basically in the

Retinex theory it is assumed that the HVS performs some random scan of the observed scene and in so doing calculates the ratio of radiant energy (reflected light) on both sides of the boundary between different objects in the scene. If illuminance is slowly changing then since the chosen points are on either side of a boundary they are close together and the illuminance at both points will be approximately the same. The ratio of reflected light will thus tend to the ratio of reflectances. In this way the reflectance at each point can be calculated. This theory has been used with some success as part of a neural system for colour constancy [7]. It has not gained wide acceptance though both for computational reasons of how to scan a scene and how to differentiate between objects and for physiological reasons amongst them being how the HVS could perform this scan, recognition and reflectance calculation in real time.

In this report we do not consider colour constancy. We are rather concerned with how detail and apparent sharpness of a colour image can be improved by manipulating colour.

A digital colour image is generally represented as three planes such that at each spatial coordinate intensities of independent primaries such as red, green and blue are known. Owing to this complex three dimensional nature of colour and to the fact that colour perception is not well understood, little research has been done on colour image detail enhancement. In fact as Hunt and Kubler note [8], none of the major textbooks on image processing contain any reference to processing of images with a higher degree of spectral complexity than monochrome ([9–11]). There seems to be an unspoken agreement that "we can process each of the three monochrome images separately and combine the results" [12].

Hunt and Kubler themselves approached the problem of image restoration and deblurring for multispectral images [8] by developing a technique which is based on optimal (Wiener) methods. A further approach to colour image enhancement has been to consider colour spaces other than RGB. One of the more natural ways of describing a colour is in terms of what colour it is (hue), how deep the colour is (saturation) and how bright the colour is (intensity). Thus some effort has been made to enhance in this HSI (hue, intensity, saturation) space. Apart from simple global saturation and brightness increases other manipulations have been attempted. Lehar and Stevens have presented a fast interactive method for saturation stretching [13]. This is done by encoding the image colour data using a peano curve into a much reduced set of colours stored in a lookup table. Manipulations are performed on the lookup table and so can be interactively viewed. Traditionally edge enhancement has been performed on the intensity component. Since this component is merely a weighted sum of the R,G and B components, the result of linear filtering methods used on it are not much different to applying the same techniques to each of the planes separately. An improvement to this type of approach was suggested by Strickland *et al* [14]. They noticed that the saturation component contains edges which are less evident or missing in the intensity image. This change in saturation indicates detail which is not evident in the brightness image. To reveal this detail they both edge

enhance the intensity component and add to this the edges from the saturation component. The resulting image does reveal more detail than when intensity edge enhancement alone is used. Another use of HSI space for colour enhancement has been presented by Bockstein [15]. Since monochrome histogram modification techniques cannot be applied to each of the RGB planes separately without unacceptably changing the colour balance, he proposed applying histogram equalization to the intensity component alone and then equalizing the saturation component but as a function of hue. That is, for each band of hues present in the image the saturation histogram is calculated and equalized thus maintaining the same hues and only varying their saturation. Unfortunately results of this method were not presented although preliminary checks that we have performed seem to reveal that colour balance is changed and that the method is more suited to enhancement for information extraction than for visual quality improvement. The method is also computationally intensive.

The shortcomings of all these techniques are due to a lack of a quantitive definition of subjective image quality. The only way to circumvent this lack is to base the processing method on knowledge of the HVS.

Very simply put, the latest generally accepted way in which the early HVS functions is that light, $I(\lambda)$, enters the eye and is absorbed in the retina by three cone types. The responses of these cone types are then transformed to a so called opponent colour stage where cells are excited by certain colours and inhibited by the opponent colour. The opponent processes consist of two chromatic channels, red-green and blue-yellow, and an achromatic or black-white channel.

Faugeras developed an image processing model based on this visual model [16,17]. A block diagram of his model is shown in figure 1.1. The first stage represented by $S_R(\lambda)$, $S_G(\lambda)$ and $S_B(\lambda)$ is a linear transformation of the cones and is followed by a nonlinear, logarithmic operation representing the cones non-linear response. These processed stimulus values R^*, G^* and B^* are then followed by the opponent stage where the chromatic and achromatic components are separated. The outputs are the luminosity variable L (black-white) and the two chromatic variables C_1 (red-green) and C_2 (yellow-blue).

A flaw in this model is that the transform to the opponent space is not known and so Faugeras had to somehow estimate it. Using the spectral responses of the cones, Faugeras employed the CIE Standard Observer relative luminosity efficiency function $V(\lambda)$ to compute α, β and γ under achromatic conditions. Thus the model is only entirely correct for achromatic images.

Using this model and Fourier techniques Faugeras showed how dynamic range can be improved and how a measure of colour constancy can be obtained.

In the rest of this report we develop a similar opponent colour model but we use new knowledge of the cone response stage and most importantly we base the opponent transform on the latest insight into how this is performed in the HVS. The enhancement we propose to

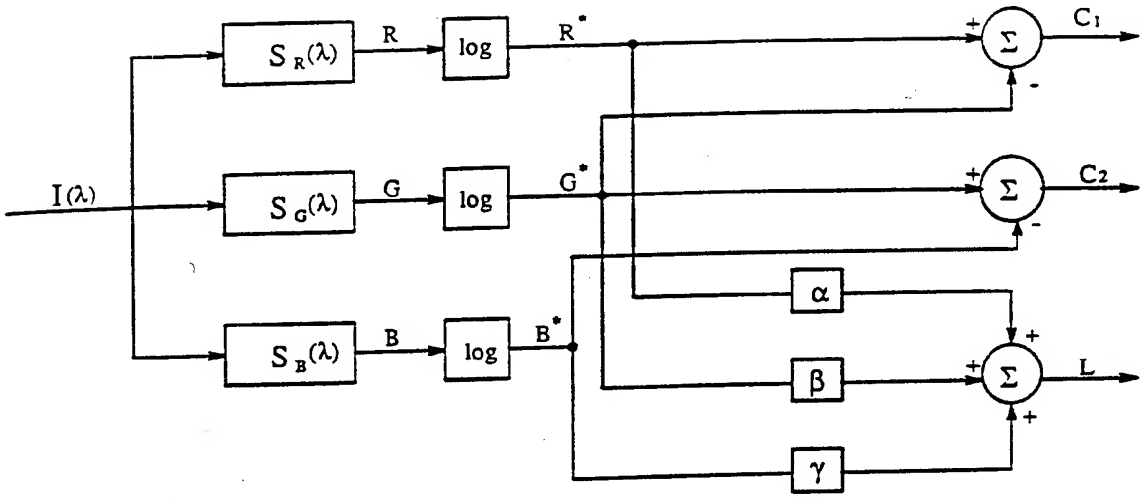


Figure 1.1: Block diagram of Faugeras image processing model reproduced from [17].
איור 1.1: דיאגרמת מלבנים של המודל לעיבוד תמונות של פוג'דה (Faugeras) מועתק מ [17].

use in conjunction with this model is also based on new knowledge of the HVS rather than on Fourier techniques. In particular, new knowledge of the existence of opponent colour centre-surround retinal receptive fields has led us to develop a colour edge enhancement technique.

In chapter 2 we give a brief review of biological colour vision which leads on to a discussion of the role of the opponent colour transform. In chapter 3 we incorporate the knowledge of the HVS discussed in chapter 2 into an opponent colour computational model of the early HVS. We first describe how the cone responses are obtained from the digital image and then how the opponent transform is calculated using a Karhunen-Loéve transform. In theory the transform is dynamically calculated for each scene but we show that because of the statistical properties of certain classes of image, we are able to derive a scene independent transform that can be used for any image in that class. We also illustrate some of the properties of the produced opponent channels and we relate the responses in these channels to the more familiar perceptual quantities of hue, saturation and intensity.

In chapter 4 we develop and describe an enhancement method that may be used on each of the opponent channels. This begins with a discussion of the role of lateral inhibition in the retina and of colour Mach bands. In this chapter we also briefly discuss a measure of colour image enhancement that is used to compare results. We then apply the enhancement technique to a few images and illustrate the results. Finally, we discuss some of the effects of noise on the model channels and show how chromatic enhancement can be performed even in the presence of noise.

Chapter 2

Background to Colour Vision

Colour is a phenomenon in which we visually perceive some function of the spectral content of the radiant energy emanating from an object. Since radiant energy is generally a complex function of wavelength the visual process is a filtering of the input signal to result in a coded version characterized by a colour name. The details of this process are not well understood although the workings of the early HVS have been quite thoroughly investigated [18, 19].

In this section we will discuss some of the findings of these investigations into the early HVS, in particular those that will help understand the computational model which we develop in the next chapter.

2.1 Trichromacy

In about 1800 Thomas Young suggested that physiological colour processing in man must be restricted to the involvement of three independent variables. In 1866 Helmholtz published the hypothesis that there exist three different channels with three different filters and even proposed the spectral response curves of these filters. Later Maxwell provided quantitive confirmation of the so called Young-Helmholtz trichromatic theory. More recent evidence confirming the trichromatic theory includes that of Buchsbaum and Gottschalk who showed that practical colours are either suitably frequency limited to be represented by three samples or have suitably frequency limited metamers [20].

The three independent colour processes in the retina are obtained from the independent processes of the three retinal cone types. These in turn are due to the three different photopigments found in each of the cone types. By measuring the spectral sensitivity of these pigments, the cone spectral response curves have been estimated. These curves have commonly been called Red, Green and Blue but this is misleading since these names neither describe the colour associated with the sensitivity curve nor of the photopigment. In this report we prefer to use the names long (l), medium (m) and short (s) wavelength responses. An example of the cone response curves are shown in figure 2.1. In appendix A we briefly

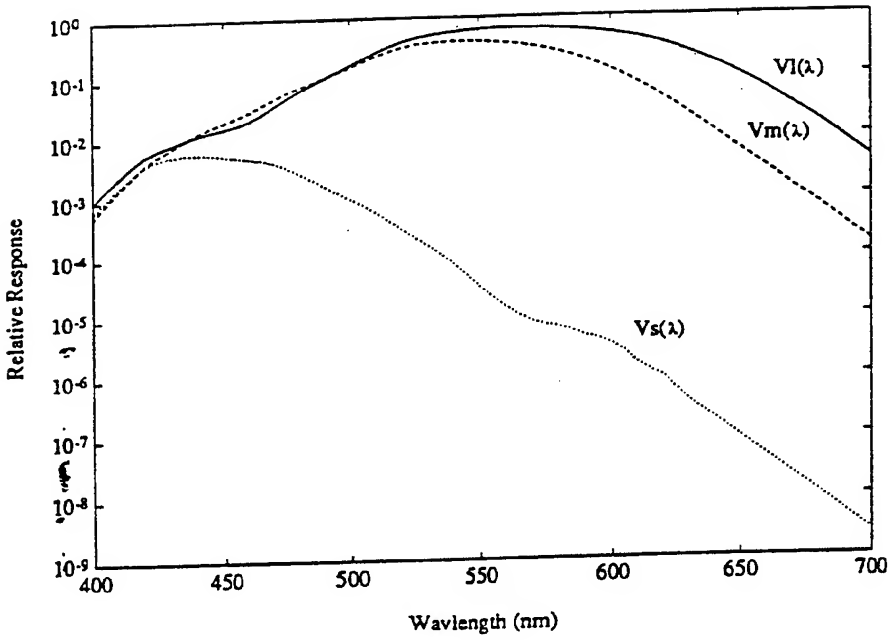


Figure 2.1: Vos-Walraven cone spectral sensitivity functions $V_l(\lambda)$, $V_m(\lambda)$, and $V_s(\lambda)$ [18]

איור 2.1: עקומי רגישות של המדוחים לאורך גל של Vos-Walraven [18]

discuss how the trichromatic theory is used for colour naming and in digital images.

2.2 Colour Opponency

There are many colour phenomena which the the simple trichromatic theory cannot explain. For example above 600 nm the short wavelength cones are ineffective: We therefore would expect that stimulation above this wavelength would appear greenish red when in fact it appears yellowish red or orange.

Some of these problems were recognized in 1878 by Ewald Hering who proposed an alternate theory for the perception of colour called the opponent-colour theory [21]. Hering proposed three mechanisms that mediate in colour vision: one that accounts for the perception of red-green, the second accounts for blue-yellow and the third for black and white distinctions. This theory was not widely accepted particularly with the physiological discovery of the three cone types.

Relatively recently at least four varieties of so-called spectrally opponent cells have been found in the LGN (lateral geniculate cortex) [16,19]. These cells have shown excitation to some wavelengths and inhibition to others when the eye was stimulated by flashes of monochro-

matic lights. Moreover this type of response was maintained over wide intensity ranges thus indicating that the cells responded to colour rather than luminance. The cells whose responses showed maximum excitation around 500 nm and maximum inhibition around 630 nm were called green excitatory, red inhibitory cells (+G-R). Other cells showing rough mirror image responses were called red excitatory, green inhibitory (+R-G). The cells showing maximum excitation (inhibition) at 600 nm and maximum inhibition (excitation) at 440 nm were named yellow excitatory, blue inhibitory cells (+Y-B) and blue excitatory, yellow inhibitory cells (+B-Y) respectively.

In addition to these four varieties of spectrally opponent cells , two other classes of cells were found that did not give spectrally opponent responses but rather responded in the same direction (either excitation or inhibition) to lights of all wavelengths. These cells thus appear to be involved with achromatic vision and are thus named white excitatory, black inhibitory (+W-B) cells and vice versa.

More recently slightly more complex opponent cells have been found, both in the final retinal processing layer, the retinal ganglion layer as well as in an earlier retinal stage containing the horizontal cells [18, 22]. In addition to the well known intensity (black-white) opponent centre-surround cells four types of colour opponent centre-surround cells have been documented. These are red centre, green surround and its mirror image as well as a blue centre, yellow surround and its mirror image.

Thus it seems that the Young-Helmholtz trichromatic and Hering opponent theory are complementary; the first dealing with early retinal processing and the second modeling phenomena at later stages of the retina and at the LGN level.

2.3 Opponent Colour Transform

Although we know that radiant energy is mapped into three independent variables by the cone responses and later into three opponent channels, the transform from the cone trichromatic responses to the opponent colour responses of the retinal ganglion cells is still not clear.

Many psychophysically based models of this opponent process have been proposed though little physiological backing has been found [16–18, 23].

Recently though, by considering the HVS as an efficient information processing system, some surprising results have emerged. What we mean by "efficient" is that redundant information is reduced and energy compressed such that the channel capacity required to transmit the information at a given level of reliability is minimized. Reasons for considering this approach in the HVS are that natural visual scenes exhibit both spatial and chromatic regularities, making the retinal image contain much redundancy. The ratio of photoreceptors to optic nerve fibres is at best 10 : 1 if we only consider cones and about 150 : 1 if we take rods into account as well [22]. Thus if the initial photoreceptor responses were left unaltered, the lim-

ited channel capacity of the nerve fibres leading to higher stages of the HVS would be unable to cope. This reasoning has led to the hypothesis that the purpose of image transformation in the early visual system is to reduce correlation and compress information prior to neural transmission.

Based on this hypothesis, Srinivasan *et al* [24] accounted for the spatial profile of center-surround retinal ganglion cell receptive fields using a linear predictive filter. They showed that the spatially antagonistic centre-surround structure of these receptive fields decorrelate and suppress an image signal. As such they can be viewed as method for removing redundant information arising from the high *spatial* correlation of natural scenes.

Using similar reasoning Tsukamoto *et al* [25] demonstrated that a gaussian spatial profile of centre-surround receptive fields is optimal in terms of spatial correlation reduction. The profile of such retinal receptive fields are indeed a close approximation to being gaussian.

Apart from the chromatic regularities of natural scenes, additional chromatic correlation is due to the highly overlapping nature of the cone chromatic response functions as can be seen in figure 2.1. In a study of this, Buchsbaum and Gottschalk [26] used an eigenvector analysis to investigate the role of opponent type processing in colour vision as well as the relation between opponent colour transforms and the initial cone responses. They showed that redundancy reduction and information compression is achieved by a transformation to an achromatic and two opponent chromatic channels. Thus the opponent transform can be viewed as method for removing redundant information arising from the high *chromatic* correlation of natural scenes.

Derrico and Buchsbaum [27] have combined these two approaches of spatial and chromatic correlation reduction into a computational model of image *coding* in early vision. The first stage of their model consists of performing an eigenvector transformation on the original red and green planes to reduce chromatic correlation. They then use linear predictive coding (LPC) on the achromatic channel alone to reduce spatial correlation. By applying this coding technique to only the red and green planes of a small group of images they transform an image into two channels which are known to have correlates in visual system psychophysics.

What we can conclude is that treating the HVS as an information processing system is a good approach. Also as regards the opponent transform, it seems that by transforming the cone responses in such a way that redundancy is reduced and energy is compressed, we should be approximating very well the operation that the HVS performs when it transforms from photoreceptor to opponent responses.

2.4 Contrast Sensitivity

Although the transformation to the opponent channels has been quite unclear to this date, some of the properties of these channels have been psychophysically investigated. In partic-

ular we are interested in the different way in which luminance and chrominance frequency information is perceived. Various models for luminance and chrominance contrast sensitivity functions (CSF) have been proposed, for example [28]. All studies agree on one point. That is, the HVS is more sensitive to high frequency luminance contrast than it is to high frequency chromatic contrast. In fact above a spatial frequency of about 3 cycles/degree the HVS is insensitive to chromatic contrast while it is still quite sensitive to luminance contrast. This variance in sensitivity is particularly relevant when high frequency enhancement is considered. It should be noted that the blue-yellow chromatic channel is even less sensitive than the red-green channel to chromatic contrast. An example of typical luminance and chrominance contrast sensitivity functions are shown in figure 2.2.

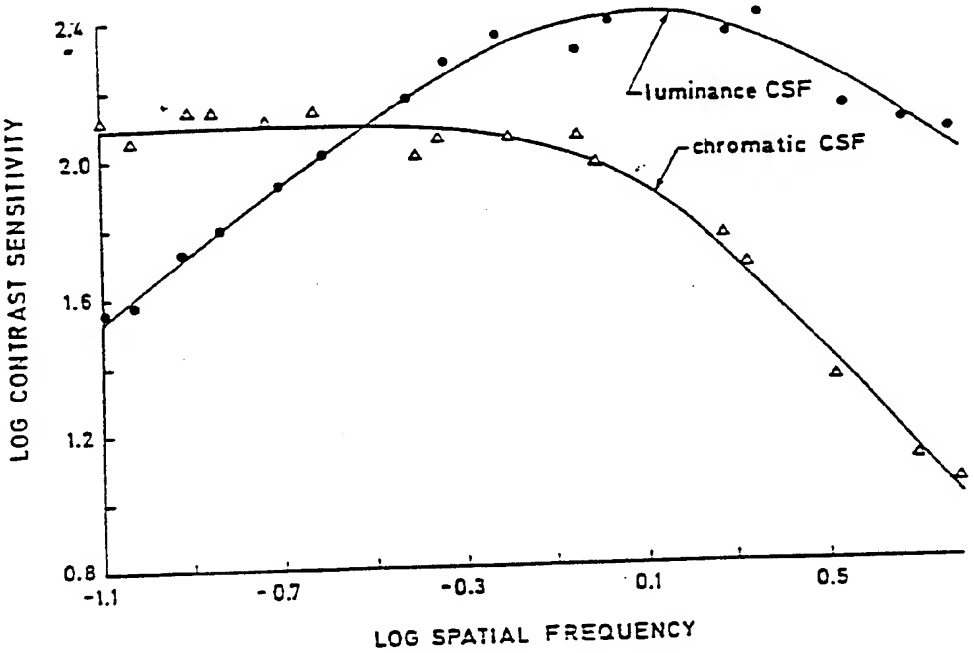


Figure 2.2: Typical psychophysical chromatic and luminance CSF's reproduced from [28]. The chromatic CSF is only shown for the red-green channel. The bandwidth of the blue-yellow channel is less than the red-green channel.

איור 2.2: עקומי ניגודיות טיפוסים של ערוצי הבהיקות והצבע במערכת הראיה מושתק מ- [28] . עקום הניגודיות של ערוץ הצבע מוצג רק לערוץ האדום-ירוק. רוחב סרט של ערוץ הכחול-צהוב יותר קטן מרוחב הסרט של ערוץ האדום-ירוק.

Chapter 3

Development of Colour Image Enhancement Computational Model

3.1 Transformation of RGB to Cone Primaries

As described in section 2.1 there are three cone types involved in human colour vision each of which contains a pigment of different spectral absorption characteristics. Based on the trichromatic theory, colours may be thought as being represented at the first stage of the retina by the tristimulus values derived from these absorption curves. These curves are termed $V_l(\lambda)$, $V_m(\lambda)$, and $V_s(\lambda)$, where l,m,s stand for the long, medium and short wavelength responses as previously discussed (section 2.1). Colours in a digital image, on the other hand, are comprised of the tristimulus values derived from the spectral functions of the camera sensors. These are commonly called Red, Green and Blue owing to the location along the wavelength axis of the respective peaks of these spectral functions. The location of these peaks imply for example that an image with RGB=(0,0,1) will look entirely blue.

There are two reasons why we require that digital colour images be represented by the retinal tristimulus values:

- We wish to use the model to enhance colour images for the benefit of a human viewer. Therefore we would like to exploit characteristics of the HVS. For example in achromatic images we increase luminance detail by enhancing luminance edges. This is done on the achromatic image itself since it represents luminance. In the same way we would like to perform enhancements on colour images in the HVS colour space.
- If we assume that the HVS transforms to the opponent-colour space using some type of information transmission analysis on the cone responses, then performing such an analysis on other tristimuli will not guarantee that the resultant space is similar to the

HVS opponent space.

Hence the first stage of the model must be to transform the image represented by the RGB set of primaries to the HVS set of primaries.

It is known from Grassmans' laws that the HVS spectral absorption curves must be a linear combination of the $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ curves, derived by the CIE from colour matching experiments [19]. However the particular linear combination is not known. Thus in the literature there exist a number of "fundamental" systems [18]. The two most widely accepted sets of fundamentals are those of Smith and Pokorny and of Vos and Walraven [18]. These two sets of spectral absorption curves are very similar. The main difference is the relative response of the short wavelength response function $V_s(\lambda)$. Buchsbaum has shown that if we form opponent channels following information transmission efficiency considerations then for a large range of relative responses of $V_s(\lambda)$, the resulting channels are independent of the actual response [26]. Since this is the basis of our model, the choice between these two sets of curves is not critical and we have arbitrarily chosen to use the Vos-Walraven set. These are illustrated in figure 2.1.

By the trichromatic theory, if the fundamentals in a set of primary colours are linearly independent, then it is possible to derive a linear transformation from a colour represented by one set of primaries to a representation by another set [18,29]. We wish to find the 3×3 transformation matrix, T_{RV}, that transforms RGB tristimuli to $V_l V_m V_s$. That is,

$$\begin{bmatrix} V_l \\ V_m \\ V_s \end{bmatrix} = T_{RV} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{3.1}$$

The transformations from RGB to XYZ, T_{RX} [30], as well as from $V_l V_m V_s$ to $\hat{X}\hat{Y}\hat{Z}$, $T_{V\hat{X}}$ [18, 30] are linear. The transformation from $\hat{X}\hat{Y}\hat{Z}$ to XYZ which may be implemented using the Vos equations [31], is a non-linear function of the luminance and is constant only if Y is unchanging. Rather than derive the transformation for each pixel in the colour image we have assumed that for image processing purposes, the $\hat{X}\hat{Y}\hat{Z}$ and XYZ values are sufficiently close that this transformation is not necessary. This approach is verified by Guth who has used the same approximation in his latest model of colour vision [23]. We have adopted this simplification since it accelerates the transformation. This is important if the model is to be used for real-time applications.

With this simplification in mind, using matrix algebra we get,

$$T_{RV} = (T_{V\hat{X}})^{-1} \cdot T_{RX} \tag{3.2}$$

which results in,

$$T_{RV} = \begin{bmatrix} 2.5653e-01 & 3.8702e-01 & 7.9427e-01 \\ 4.2417e-01 & 2.5035e-01 & 1.7708e-01 \\ 3.6682e-08 & 2.4582e-04 & 4.1504e-03 \end{bmatrix} \quad (3.3)$$

Thus at each spatial location in the image we transform the RGB tristimuli to $V_l V_m V_s$ using the colour transform, T_{RV}.

3.2 Separation of Chromatic and Achromatic Information

The discussion of section 2.3 has led us to believe that the approach of considering the HVS as an efficient information processing system is a good one and that the most natural transformation for colour images is one that will result in the best correlation reduction and energy compression. The optimum transformation in terms of these requirements is the Karhunen-Loéve transform (KLT). The KLT is optimal since it can be shown that the error between the original image and an image reconstructed from a properly selected subset of KLT components is the smallest amongst all linear image transforms [9,10,32-36]. Therefore the next step after transforming from RGB to $V_l V_m V_s$, is to apply the KLT to decorrelate the V_l, V_m and V_s cone responses at each spatial coordinate. In the next two sections we explain how the Karhunen-Loéve transform is calculated.

3.2.1 Discrete Karhunen-Loéve Transform

This transform also known as the principal component, eigenvector or Hotelling transform can be formulated as follows:

If $\{X(n), 1 \leq n \leq N\}$ is a random sequence with autocorrelation matrix R and Φ is an N by N unitary[1] matrix that reduces R to diagonal form then the transformed vector

$$Z = \Phi X \quad (3.4)$$

is called the discrete Karhunen-Loéve transform of X. It can be shown that the rows of Φ are the normalized eigenvectors of R [9,10,32-36] and that these eigenvectors are also orthonormal [37]. Therefore the elements of the sequence $Z(n)$ are orthogonal. It can also be shown that if Φ is made up of the normalized eigenvectors of the covariance matrix C rather than of the autocorrelation matrix R, then the sequence $Z(n)$ will also be decorrelated. The covariance matrix is simply the correlation matrix of a zero mean process or alternatively, the correlation matrix of a non-zero mean process that has been centralized. In other words, $R(n,m) = E[x_n x_m]$ and $C(n,m) = E[(x_n - m_x)(x_m - m_x]$ where m_x is the mean value of x.

[1] A unitary matrix is one such that its inverse equals its conjugate transposed. i.e. $A^{-1} = (A^*)^T$.

3.2.2 Multispectral Discrete Karhunen-Loéve Transform

We now adapt the discrete Karhunen-Loéve transform to the multispectral case. In a multispectral image, a scene consists of M spectral planes. In the case of RGB, $M = 3$.

Let $\mathbf{X}(k,l)$ represent the intensity vector whose M elements are the values of the pixels at location (k,l) in the set of spectral planes. That is,

$$\mathbf{X}(k,l) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix}$$

where

$$\begin{aligned} k &= 1,2,\ldots,L \\ l &= 1,2,\ldots,W \\ x_i &= \text{intensity at (k,l) for plane i, and} \end{aligned}$$

L and W are the number of pixels per row and column respectively, of each spectral plane.

$\mathbf{X}(k,l)$ is the random vector whose components we wish to decorrelate. Since there are $L \cdot W$ such vectors, under the assumption of stationarity and ergodicity we may estimate the covariance matrix as shown below. Note that the assumption of stationarity is not in general entirely true because of the changing spatial characteristics of natural images.

The covariance matrix of dimension $M \times M$, is defined as,

$$\mathbf{C_x} = E[(\mathbf{X} - \mathbf{m_x})(\mathbf{X} - \mathbf{m_x})^t]$$

where

$$\mathbf{m_x} = E[\mathbf{x}]$$

It can be approximated under the above assumptions, from the $L \cdot W$ samples [9,38], by,

$$\mathbf{C_x} \approx \frac{1}{L \cdot W} \sum_{k=1}^{L} \sum_{l=1}^{W} \{[\mathbf{X}(k,l) - \mathbf{m_x}][\mathbf{X}(k,l) - \mathbf{m_x}]^t\} \tag{3.5}$$

and

$$\mathbf{m_x} \approx \frac{1}{L \cdot W} \sum_{k=1}^{L} \sum_{l=1}^{W} \mathbf{X}(k,l) \tag{3.6}$$

Using the image data in expression 3.5 then results in the M $\times$ M covariance matrix, $\mathbf{C_x}$. Since M is usually small, it is a relatively simple matter to find the eigenvectors and

eigenvalues of C_x. If we denote the normalized eigenvectors as $\bar{e}_n$ and the corresponding eigenvalues by λ_n then the transform matrix Φ is:

$$\Phi = \begin{bmatrix} \bar{e}_{11} & \bar{e}_{12} & \cdots & \bar{e}_{1M} \\ \vdots & \vdots & \vdots & \vdots \\ \bar{e}_{n1} & \bar{e}_{n2} & \cdots & \bar{e}_{nM} \end{bmatrix}$$

where e_{ij} is the j'th element of the i'th normalized eigenvector.

Eigenvectors are normalized in the usual way,

$$\bar{e_n} = \frac{e_n}{|e_n|}$$

That is, the rows of Φ are the normalized eigenvectors of C_x. By ordering the rows of Φ such that the corresponding λ_n are in descending order, the *energies* of the resulting transformed images will also be in descending order.

It can easily be shown that the covariance C_z, of the transformation $Z(k,l) = \Phi X(k,l)$ is equal to $\Phi C_x \Phi^{-1}$. It can also be shown that:

$$C_z = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \lambda_M \end{bmatrix} \tag{3.7}$$

where $\lambda_1 > \lambda_2 > \ldots > \lambda_M$ are the eigenvalues of C_x. Thus, the transformation produces statistically decorrelated images, at least in the sense of having a diagonal covariance matrix.

To summarize, we transform to the opponent space by using expression 3.5 to estimate the covariance matrix of the cone response data. We then calculate the eigenvectors of this matrix, and arrange the normalized eigenvectors such that the respective eigenvalues are in decreasing order. This then gives us the transformation, Φ, to the opponent space.

We then use Φ to transform each colour vector at every spatial location in the image. This results in a rotation of the V_l, V_m and V_s axes, forming a new set of orthogonal and uncorrelated axes which we label K_1, K_2 and K_3 that span the colour space. Unlike a previous study [26], the resulting channels have a spatial structure whose nature we exploit not for coding purposes but for enhancement.

3.3 The Enhancement Model and Its Relation to Other Models

Now that we have represented the digital colour image in a space closely allied to that the HVS uses, we perform enhancement on each of the K_i channels. The actual type of enhancement we consider is explained elsewhere (section 4.2). To view the results we have to inverse transform back to RGB. It is clear that $(T_{RV})^{-1}$ exists but we also have to be sure that Φ^{-1} exists. Fortunately, because C_x is real and symmetric, and therefore its eigenvectors are orthonormal [9], we have that:

$$\Phi^{-1} = \Phi^t \qquad (3.8)$$

The inverse thus always exists and inverse transformation is always possible.
We are now in a position to draw a block diagram of the model. This is shown in figure 3.1.

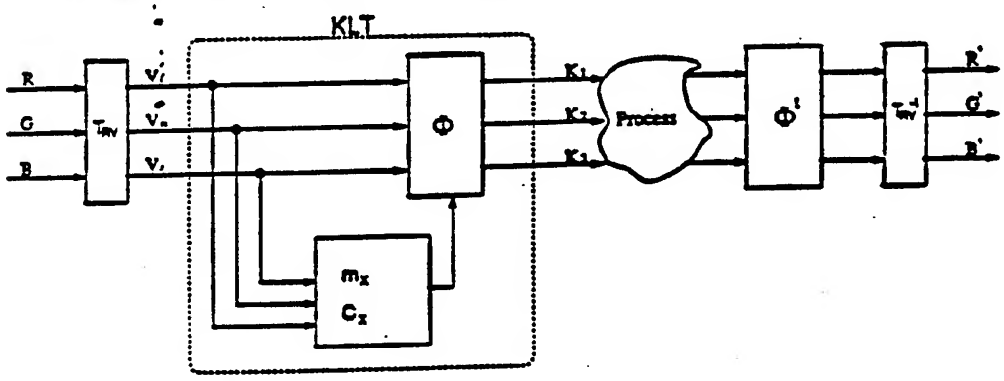


Figure 3.1: Block diagram of the generalized computational model

איור 3.1: דיאגרמת מלבנים של תהליך המידול הכללי.

Since C_x is real and positive it will always have *only* one eigenvector whose entries are of like sign [26]. This means that only one of the transformed channels generated by Φ can be an all positive combination of V_l, V_m and V_s. The other two channels must therefore be antagonistic. The all positive channel clearly contains the greatest proportion of the image energy and is associated with the largest eigenvalue. This channel is labelled K_1. The two other antagonistic channels are labelled K_2 and K_3.

We show in section 3.4 that the wavelength response function of the K_1 channel approximates well the HVS luminous efficiency function. We therefore call this the achromatic channel. We also show that the wavelength response functions of the K_2 and K_3 channels are similar to the response functions of the opponent red-green and blue-yellow channels psychophysically found in the HVS [18,19]. We therefore classify our model as an opponent colour

model along the lines of the theory first proposed by Hering [21]. Many attempts have been made to quantify his theory, the most recent of which being by Guth [23]. Faugeras [16,17] has been the only one however to attempt to use an opponent colour model for image enhancement purposes.

The main difference between this model and Faugeras' is that we dynamically transform to the opponent channels for each scene based on an emulation of the behavior of the HVS. Faugeras transforms using a constant transformation that he derived from psychophysical experiments. In section 3.4 we describe an experiment that shows that for scenes belonging to a similar class our model may be reduced to a constant transform. We may thus consider it to be a more generalized image processing opponent colour model.

Another difference between our model and Faugeras' is that he transforms the logarithm of the cone responses. Thus for black and white images, his model reduces to Stockham's homomorphic model for achromatic vision. Stockham proposed a multiplicative homomorphic model for processing achromatic images that matched both the structure of images and the HVS [39]. The multiplicative process of image formation where an achromatic image $A(x,y)$, is the product of an illuminance function $i(x,y,\lambda)$ and a reflectance function $r(x,y,\lambda)$ can be undone by a logarithmic operation. This is because the logarithm is a homomorphic function from multiplication to addition. Hence it maps the image product into the sum of two components which can then be separated by linear filtering since in general the frequency content of $i(x,y,\lambda)$ and $r(x,y,\lambda)$ is distinct. Stockham's model also matches the HVS since the cone responses are non linear for very high and very low levels of illumination. His model is especially useful for improving dynamic range.

Faugeras argued that as the frequency content of the luminance component that is associated with the illuminant is different to the reflectance, then so should the frequency content of the colour components, which are associated with the colour of the illuminant, be different. Faugeras thus also used a logarithmic non-linearity to achieve dynamic range enhancement in his luminance channel and a measure of colour constancy in the chromatic channels. Because of the three dimensional nature of colour vision however, a colour image cannot be strictly considered as simply a product of illuminance and reflectance functions. Colour image formation is a multiplicative process that yields the product of $i(x,y,\lambda)$ and $r(x,y,\lambda)$ but before the visual system (HVS or camera) processes this product, the light is linearly absorbed by three sensors or cone types as expressed by equations A.1– A.3. We no longer have a logarithm of a product but the logarithm of a sum of products. Although it has been shown that under certain circumstances it may be possible to approximate the image tristimulus values as products of illuminance and reflectance tristimuli [40] this is in general not true. Also, while it is true that the cone response is approximately logarithmic, we are not sure what other non-linearities there are prior to transformation. For these reasons, we have elected not to use a homomorphic model and not to include the logarithmic non-linearity at this stage.

As a final comment, we know that the problem of colour constancy is far more complex than Faugeras' model suggests. We also know that if desired, dynamic range can be improved by post-processing on R', G' and B' using a logarithmic scaling method as suggested by Pratt [34]. Thus, in terms of image processing, taking the logarithm of the cone responses is not of much benefit.

3.4 Scene Independent Transform

Up till now, all opponent colour models have been based on a constant transform from cone responses to chromatic channels [16–18,23]. This transformation has been derived by fitting model parameters to various sets of psychophysical data. This is an acceptable approach if the models so derived are to be used only for predicting psychophysical responses. When the visual stimulus is more complex than sinusoidal gratings or coloured patches, we cannot be certain that these models accurately reflect the behavior of the HVS. However, it is reasonable for a number of reasons to assume that the opponent transform is constant. Firstly, the transform is from photoreceptors to ganglion cell responses and is thus performed entirely in the retina and secondly, the transform is performed in real time. This is not to say that the colour channels of the HVS are non-adaptive but that adaption to a particular scene and illuminant is gradual. Adaption aids in our ability to recognize colours under many different illumination conditions (colour constancy) but we assume that under any one particular condition of illumination the colour transform is constant. Adaption is also assumed to be performed later on in the HVS and not in the retina.

In what follows we show how our model, which is able to take into account the complex nature of scenes, because of the use of the Karhunen-Loéve transform, can also be used to derive a constant transform for complex scenes of a similar type.

Based on the assumption that man has adapted to his environment, it is reasonable to assume that the HVS has optimally adapted to scenes which it most often encounters. Such colour scenes are natural daylight type images. If this is the case then the opponent colour transform should maximally decorrelate and compress such images. If we consider the converse of this, then maximal compression and decorrelation of natural images should yield the opponent transform present in the HVS.

In this section we have *experimentally* investigated this idea using a real set of natural images. Seven colour images of various natural outdoor scenes, similar to those shown in section 4.4, were used to derive the constant transform. They were taken with different cameras and under different daylight illumination conditions. In this way we hoped to obtain as good a representative ensemble of images of this type as possible. It must be emphasized though that although the set of images consists of a large number of triplets of red, green and blue pixels, this in itself is no proof that the images we chose are representative of the entire

ensemble of natural images. Our results do seem to suggest however that these images are a reasonably good representation.

To derive the constant transform, we first have to estimate the chromatic covariance matrix of this set of images. Since spatial correlations are of no consequence for this, one method is to simply form a single super-set consisting of all the available colour triplets. This super-set could then be used to estimate the colour covariance. The problem with this method is that owing to the immense amount of data ($7 \times 512 \times 512 \times 3 = 5.25\ Mb$ for 7 images of resolution 512×512), it is not feasible to use directly expressions 3.5 and 3.6. For this reason, we derived an estimate of the covariance of a set of data based solely on the means and covariances of each image in the set. Details of this derivation may be found in Appendix B. The derivation resulted in,

$$\mathbf{C} = \frac{1}{N}\left[\sum_{n=1}^{N}\left(\mathbf{C}_n + \mathbf{m}_n\mathbf{m}_n^t\right) - \frac{1}{N}\left(\sum_{n=1}^{N}\mathbf{m}_n\right)\left(\sum_{n=1}^{N}\mathbf{m}_n^t\right)\right] \tag{3.9}$$

where

n = $1, 2, \ldots, N$
N = number of images in the set.
$\mathbf{C}_n$ = covariance matrix of the n'th image.
$\mathbf{m}_n$ = mean vector of the n'th image.
$\mathbf{C}$ = covariance matrix of the set of images.

Applying equation 3.9 to the set of images and calculating the normalized eigenvectors resulted in,

$$\Phi_c = \begin{bmatrix} 0.9517 & 0.3070 & 0.0032 \\ 0.3068 & -0.9516 & 0.0193 \\ -0.0089 & 0.0174 & 0.9998 \end{bmatrix} \tag{3.10}$$

Buchsbaum *theoretically* verified that maximal decorrelation and compression predict an opponent colour transform similar to that psychophysically found in the HVS [26]. He did this by assuming that the image, as a function of wavelength, that reaches our eye belongs to an ensemble of images which have a broadband Fourier frequency power spectrum. This is true only for images made up of monochromatic signals such as those used in visual psychophysics. The constant transform Φ_c which we have found using natural images, not necessarily made up of monochromatic colours, is similar to the transformation matrix that Buchsbaum theoretically calculated. This result reinforces Buchsbaums conclusion that the goal of the opponent transform in the retina is to achieve redundancy reduction and information compression. It also shows that the HVS seems to be optimally adapted to natural scenes.

To illustrate that our model is indeed an opponent colour model and that the colour space we transform to is similar to the HVS colour space, we use Φ_e to transform the Vos-Walraven curves shown in figure 2.1. This results in the chromatic response functions shown in figure 3.2. In figure 3.2 (a) we can see the close agreement between the model predicted K_1 channel and the relative luminous efficiency curves derived by Judd [18]. Also, the K_2 and K_3 chromatic response functions shown in figure 3.2 (b) and figure 3.2 (c) respectively, are similar to the red-green and blue-yellow opponent colour channels derived by various researchers [18].

The implication of being able to derive a constant transform of a set of images in terms of image processing is that by using our model to derive constant transforms for different classes of images we can emulate a visual system that optimally adapts to the set of images with which it deals. What this means is that the expensive and time consuming pre-processing stage of deriving Φ can be performed offline. The transformation to the opponent colour space can then be performed in real time by two matrix transformations by simply identifying the scene type and using the relevant pre-determined Φ. The general model of figure 3.1 can then be reduced to the simplified model of figure 3.3.

In section 4.5 we compare the results of performing the same enhancement on an image transformed first according to the generalized model of figure 3.1 and then by the simplified model of figure 3.3 which uses the constant transform, Φ_e. Visually the difference between the two results is imperceptible.

3.5 Frequency Properties of Karhunen-Loéve Channels

Since we have used the Karhunen-Loéve transform to obtain the opponent channels and because K_1 is associated with the largest eigenvalue, K_1 will contain a significant portion of the total image energy. K_2 is associated with the next largest eigenvalue and K_3 with the next, so they will respectively contain less energy.

To illustrate this we have transformed two typical RGB colour images to the K space. These images are shown in figure 4.4 and figure 4.13. We then estimated the percentage energy in each channel by taking ratios of the eigenvalues estimated by the Karhunen-Loéve transform ($P_i = \lambda_i/(\lambda_1 + \lambda_2 + \lambda_3)$). This method can be understood since each eigenvalue λ_i is equal to the variance of K_i along eigenvector $\mathbf{e}_i$ [9], as expressed by equation 3.7. The results of these calculations are shown in table 3.1. For comparison we have also shown the amounts of energy that are present in the original red, green and blue channels. The distribution of energy in the transformed images (table 3.1), shows that practically all the energy is contained in the achromatic image (K_1) while almost no energy remains in the red-green (K_2) or blue-yellow (K_3) chromatic images. This is a result typical of the use of the Karhunen-Loéve transform on chromatic images [8,34,38,41].

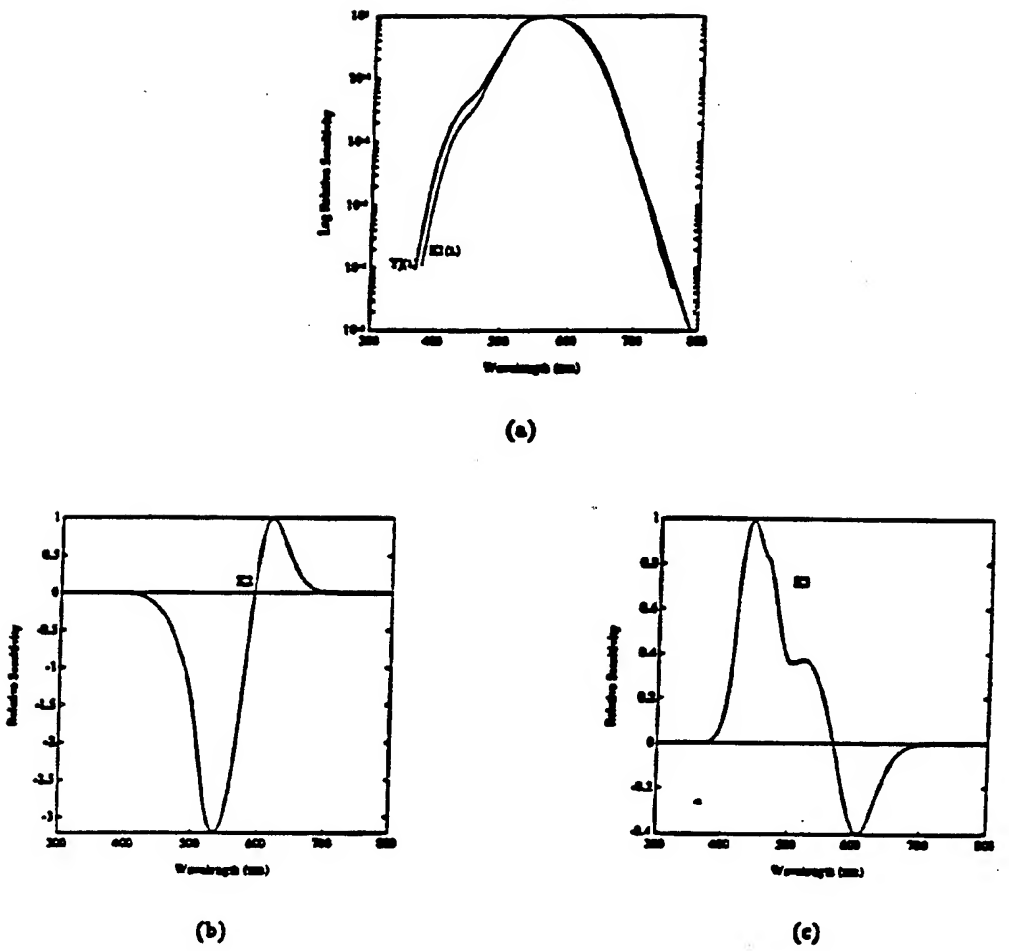


Figure 3.2: Spectral sensitivity of the achromatic (K_1) channel compared to Judds relative luminous efficiency curve is shown in (a). Shown in (b) and (c) are the spectral sensitivity curves of the red-green (K_2) and blue-yellow (K_3) channels. The K_1 curve has been shifted upwards for comparison and the K_2 and K_3 curves are normalized.

איור 3.2: (a) רגישות לאורך הגל של הערוץ האכרומטי (K_1) בהשוואה לעקומת צפילות בהיקות של ג'אד(Judd). עקומים (b) ו (c) מציגים את עקומי הרגישות של הערוץ האדום-ירוק (K_2) ושל הערוץ הכחול-צהוב (K_3). לצורך השוואה עקום (K_1) מוזז למעלה והעקומים (K_2) ו(K_3) מנורמלים.

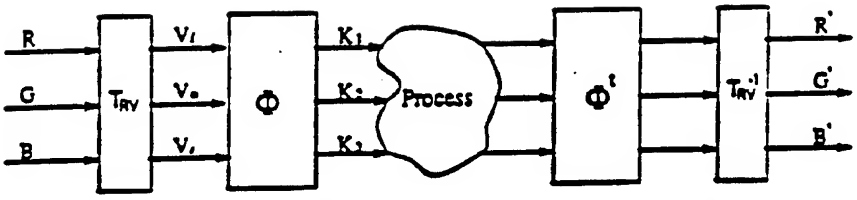


Figure 3.3: Block diagram of the simplified computational model

איור 3.3: דיאגרמת מלבנים של תהליך החישוב המפושט.

Image	R	G	B	K1	K2	K3
View	24.0%	35.7%	40.3%	99.95%	0.048%	0.002%
Ship	20.4%	37.6%	42.0%	99.94%	0.059%	0.001%

Table 3.1: Table of signal energies in the various colour planes

טבלה 3.1: טבלת אנרגיית האותות בתחומי צבע שונים.

In previous studies [8,34,38,41], where the Karhunen-Loéve transform was used for coding purposes, the *original RGB* values were Karhunen-Loéve transformed. The energy distribution in the transformed planes in these cases were of the order of $K_1 : K_2 : K_3 = 86 : 13 : 1$. The reason that we have achieved such a large concentration of energy in the K_1 channel relative to the other two channels is because we have transformed the cone tristimuli, V_l, V_m and V_s and not the original *RGB* values. Looking at figure 2.1 we see that the amplitude of $V_s(\lambda)$ relative to $V_l(\lambda)$ and $V_m(\lambda)$ is much smaller. Hence there is already a measure of information compression *prior* to transforming. If we consider the viewpoint that the Karhunen-Loéve transform is merely a repositioning of axes so as to achieve minimum variance along each axis, then it is clear that because along one axis there is already small relative variance then one of the transformed axes should be parallel or nearly parallel to the untransformed V_s axis. This is indeed the case as can be seen from the expression for Φ_t (expression 3.10). Therefore it seems that the cone response functions have also adapted in the HVS such that when followed by the opponent stage, energy compression will be maximized. This result is in keeping with the role of the HVS as an efficient information transmission system.

To understand the difference between performing the KLT in the *RGB* subspace or in the cone response space $V_lV_mV_s$ it must be noted that the *RGB* colour space of digital colour images is not an orthogonal Cartesian subspace of R^3 in the mathematical sense. That is, the RGB subspace is not spanned by mutually orthogonal vectors or in other words, the spanning vectors of the *RGB* subspace are inclined at some indeterminate angle to one another. Perceptually however Red, Green and Blue tend to be considered as independent quantities which is why in colour science they are traditionally represented on orthogonal axes

leading to tools such as the colour cube.

Since RGB is not an orthogonal subspace the transformation from RGB to the $V_l V_m V_s$ subspace cannot be viewed as a simple translation, rotation and scaling of axes. Considering T_{RV} as a linear transform of R^3 that maps a subset RGB, to another subset $V_l V_m V_s$, then it would be a simple matter to calculate the spanning vectors of $V_l V_m V_s$ from the spanning vectors of RGB [42]. Since we do not know which vectors span RGB this cannot be done. All that we can estimate at this stage is that under the standard Euclidean metric of R^n elements of subspace $V_l V_m V_s$ will be closer together along the V_s axis than they were along any dimension of the RGB subspace. This can be seen from the transformation T_{RV} (3.3). This effectively warps the pixel cloud so that variance is reduced along one dimension producing a more elliptical cloud. The process of fitting orthogonal axes to this data such that correlation is minimized along each axis (KLT) is then simplified.

Because K_1 is calculated as a weighted sum of the positive V_l, V_m and V_s tristimuli, edges and all other spatial information is preserved in this channel. We can thus think of K_1 as a spatial bandpass channel as it should contain spatial frequencies from DC up to some high frequency. K_2 and K_3 are computed as weighted differences and hence high frequency information should be reduced. This is because edges in the RGB or $V_l V_m V_s$ space are usually at the same spatial location. Subtraction of one plane from another thus reduces the amplitude of the edges and hence the amount of high frequency information. Spatially therefore we can consider them to be lowpass channels. To illustrate this we calculated the average power spectral density (PSD) for the "view" image of figure 4.4. The PSD was calculated using the classical spectral estimation method of periodogram averaging [12, 43]. The two dimensional estimate was then averaged over each radial frequency to obtain a one dimensional representation.

Shown in figure 3.4 is the PSD of the K_1 and K_2 channels. The PSD of the K_3 channel is not shown as it is almost zero for all frequencies except for those close to DC. The high frequency nature of the K_1 plane and the low frequency nature of the K_2 (and K_3) plane is evident. The low pass nature of the chromatic channels becomes significant when we consider enhancing noisy images (section 4.6).

In figure 3.5 the amplitude spectrum, which is the square root of the PSD, is shown on a logarithmic scale for all three channels. It is interesting to note that the behaviour of the K_1 channel is approximately of the form of $1/f$ on a logarithmic scale. This is in accordance with what Field has found for natural achromatic images [44]. The behaviour of the chromatic channels on the other hand is different. If we ignore the dc region of K_3 then both K_2 and K_3 fall off at a rate of approximately $1/2f$. Thus it seems that natural images not only exhibit characteristic behaviour in terms of their luminance component but also in terms of their chrominance components.

It is also interesting to compare the spatial characteristics experimentally found with the

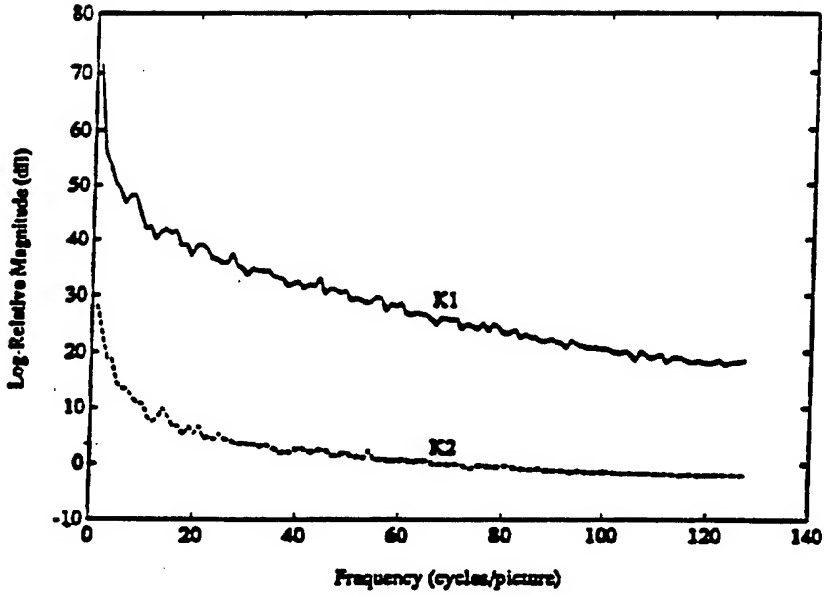


Figure 3.4: Average logarithmic power spectrum of the K_1 and K_2 planes. K_1 retains a large proportion of the total image energy at high frequencies while K_2 (and K_3) contains significant energy only for low frequencies. The frequency axis units are in cycles/picture where each picture contains 256 × 256 pixels.

איור 3.4: ספקטרום הספק לוגריתמי של שדות K_1 ו-K_2. K_1 מכיל כמות יחסית רבה של אנרגית התמונה בזמן ש-K_2 (K_3-ו) מכיל כמות משמעותית רק בתדרים נמוכים. צירי התדר מוצג ביחידות של מחזורים לתמונה. כל תמונה מכילה 256 × 256 פיקסלים.

psychophysically predicted modulation transfer function shown in figure 2.2. In the HVS, the luminance channel is indeed bandpass while the two chromatic channels are lowpass. This lends weight to the observation that the way we calculate the opponent transform results in a colour space similar to that of the HVS.

Before we leave this section it is important to note that although relative to the achromatic K_1 channel, the chromatic channels are low pass, as channels within themselves, there is still a significant amount of energy present at high frequencies.

3.6 Perceptual Interpretation of K Space

One of the more natural and familiar ways of describing a colour image is in terms of hue, saturation and intensity (HSI). We would thus like, at least roughly, to relate the responses in the K channels to the perceptual quantities of HSI. To do this, we use the constant

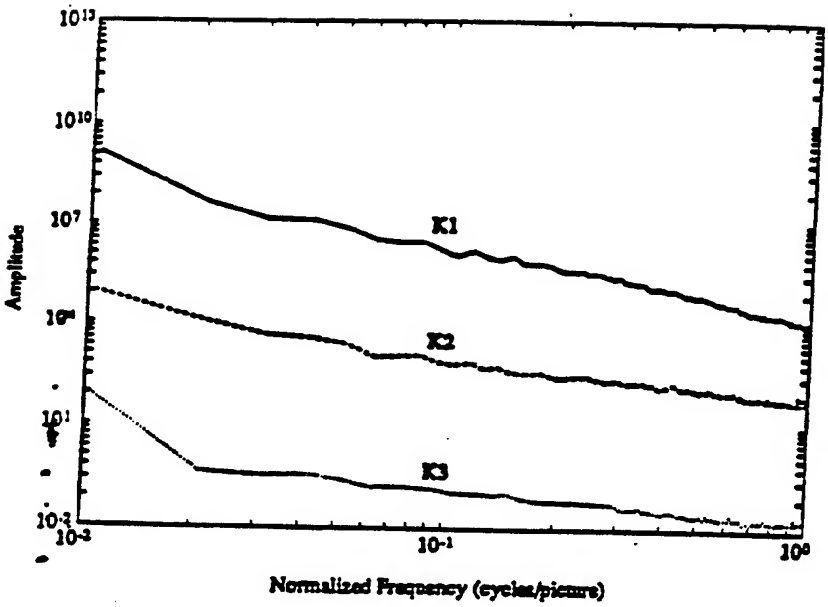


Figure 3.5: Average amplitude spectrum of the K_1, K_2 and K_3 planes represented on log-log axes. The frequency axis units are in normalized cycles/picture where each picture contains 256 x 256 pixels.

איור 3.5: ספקטרום אמפליטודה של שרוצי K_1,K_2 ו-K_3 מוצג על צירים לוגריתמים. ציר התדר מוצג ביחידות נורמליות של מחזורים לתמונה. כל תמונה מכילה 256 x 256 פיקסלים.

transform Φ_c that was derived in section 3.4. This is justifiable since Φ_c is suitable for a wide class of images that we consider the HVS has optimally adapted to and because it is the type of image that we mainly consider in this study.

From figure 3.2(a) we can see the close agreement between the model predicted K_1 channel and the relative luminous efficiency curves derived by Judd [18]. It is clear from this that K_1 contains the luminance information which is why we call it the achromatic channel. With reference to the image processing model, values of K_1 close to zero represent dark pixels while increasing positive values represent lighter and lighter greys.

The K_2 and K_3 chromatic response functions shown in figure 3.2(a) and (b) respectively, are similar to those derived psychophysically by various researchers [18]. They are clearly opponent channels and judging by the peaks along the wavelength axis, K_2 is a Red-Green opponent channel and K_3 is a Blue-Yellow opponent channel. Increasingly positive values of K_2 for example represent increasing levels of Red while increasing negative values represent increasing levels of Green. The relationship between the K space and perception of colour is shown in figure 3.6.

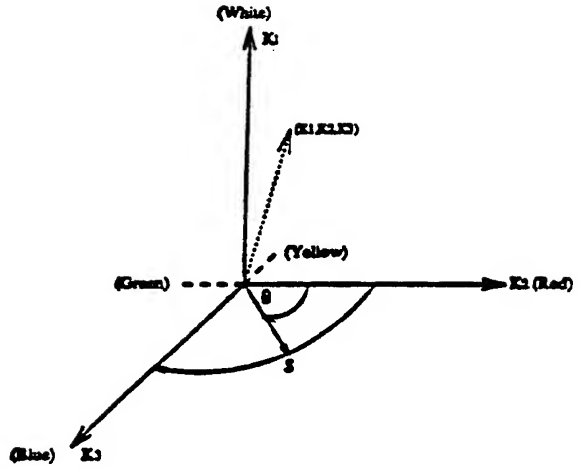


Figure 3.6: Perceptual quantities of Hue, Intensity and Saturation shown in the orthogonal opponent K space

איור 3.6: תכונות של גוון, רוויה ובהירות מוצגות במרחב המטען, האורתוגונלי, K

Intuitively, the perception of saturation is related to how "far" a colour is from being grey. For lack of a better measure of distance, we consider a Euclidean metric to calculate the distance from the $K1$ axis, giving:

$$S = \sqrt{(K_2)^2 + (K_3)^2} \tag{3.11}$$

Saturation is thus related to the sum of the activities in the two chromatic channels. According to this metric, circles in the K_2-K_3 plane represent colours of equal saturation as shown in figure 3.6.

To investigate this more thoroughly, we mapped these circles to the CIE x-y plane. We did not obtain the same curves of constant saturation as determined psychophysically [18]. This seems to suggest that the Euclidean metric is not appropriate. We then mapped ellipses whose major to minor axis ratio was in the ratio of the power present in the K_2 and K_3 channels. This resulted in curves that tended towards those psychophysically determined. What this all means in terms of image processing is that although the Euclidean metric is inappropriate, increasing distance from the K_1 axis does represent increasing saturation. The fact that the metric is not Euclidean indicates that saturation is colour dependent.

In a like manner we define hue (H) as the ratio of the activities in the K_2 and K_3 channels. Therefore geometrically, H can be thought of as the angle θ in the K_2-K_3 plane.

Chapter 4

Use of Model in Colour Enhancement

4.1 Attempted Enhancement Methods

Before discussing the method we have settled on for enhancement, we discuss some of the methods that we attempted but that were unsuccessful.

4.1.1 Amplitude Scaling

As discussed in chapter 3, the K_2 and K_3 channels are opponent red-green and blue-yellow channels. Physiologically, if we increase the neural activity in the HVS colour channels then the sensation of colour should increase. Since in our model larger absolute values of K_2 or K_3 represent increasing amounts of colour, the sensation of colour may be increased by simply increasing these absolute values. In addition, because colours are represented in the HVS in an opponent fashion, it is possible that if we increase the *difference* between the opponent colours a richer image will result. For example in the K_2 channel if we make the reds more red and the greens more green, then detail revealed by colour should become more apparent. Both an increase of total colour and of colour difference may be achieved using linear amplitude scaling. That is,

$$K_i' = m \cdot K_i + c \tag{4.1}$$

where

$$\begin{aligned} i &= 2,3 \\ m &> 1 \\ c &\geq 0 \end{aligned}$$

The constant c represents a global red-green or blue-yellow colour shift and without additional knowledge about the image is usually set equal to zero.

The effect of equation 4.1 on colour strength can be understood by substituting into equation 3.11 which results in $S' = mS$. That is, strength is increased by a factor of m. Also, since K_2 and K_3 are opponent channels and $m > 1$, multiplication by m increases the opponent colour difference.

Amplitude scaling is attractive since it is a purely point-wise operation and is thus very fast to perform. The problem with it is that unlike its use in achromatic images, here we do not know the allowable range of K_i. Therefore m cannot be calculated and must be obtained heuristically. Typical values that we used were in the range of 1.2–1.5. Note that the same m must be used for K_2 and K_3 in order to retain the correct colour relationships.

The overall effect of linear amplitude scaling is merely a sensation of global saturation or colour strength increase. There is no enhancement of detail. It may be used though in conjunction with the detail enhancement procedure described in section 4.2.

4.1.2 Histogram Methods

W_2 also showed in chapter 3 that the frequency nature of the two opponent channels is lowpass. Since they are lowpass, values in K_2 and K_3 occupy a limited range. Also because of their opponent nature, values are clustered in a small positive and negative region. As a result their histograms tend to be highly skewed with two peaks concentrated over a small range as can be seen in figure 4.1. A common technique for enhancing skewed black and white images is to use histogram modification techniques where the original image is mapped so that the histogram of the enhanced image follows some desired form [10]. The most commonly used method is histogram equalization for which the enhanced image histogram is forced to be uniform resulting in increased dynamic range.

It seems attractive to use a similar method for colour images. Owing to the three dimensional nature of colour however, naive application of histogram enhancement techniques will result in severe colour distortion. For example if a large number of pixels have a similar colour then separate equalization on each of the colour planes will result in these same pixels now spanning a range of arbitrary unrelated colours. One solution to this problem has been to perform a hue dependent equalization of the saturation component of the image [15]. Although the K channel responses can be related to the sensations of hue and saturation (section 3.6), we have rather chosen to resolve the problem directly. We did this by designing three approximate histogram modification techniques which maintain colour integrity by retaining the original pixel ordering. In other words, if in the original image planes, the value of pixel i, p_i, is less than the value of pixel j then the enhanced value of pixel i, p'_i will be less than p'_j. That is

$$p_i < p_j \Longrightarrow p'_i < p'_j$$

Thus colours will be changed in a controlled manner and colour relationships will be maintained. The three techniques are approximate equalization, modified approximate equalization and peak shifting.

Approximate Equalization

This technique is a milder form of equalization that subtly alters the original histogram in such a way that the resulting histogram of the modified channel is closer to uniform than the original. Essentially the way this is done is to first determine the mean level frequency of occurrence $\bar{f}(p)$ from the original image histogram. Then all pixels p_i which are such that $f(p_i) > \bar{f}(p)$ by some margin are reduced by that same margin and vice versa. The net result is that low count bins are increased and high count bins are reduced towards the mean level. This method therefore approximates to equalization as a more uniform histogram results. Determination of the margin is a problem that is picture dependent.

The results of this method are that saturation is increased or decreased in various regions of the picture. This is preferable to a global saturation enhancement, achieved by amplitude scaling, as slightly more detail is made evident.

Modified Approximate Equalization

This method is essentially the same as approximate equalization except that additional bins are added to the lower and upper regions to extend the colour range. This has a counterpart in black and white terms to dynamic range enhancement where the enhanced black and white image is forced to span the range [0,255] as uniformly as possible. In the case of our model, the desired range of the colour channels is unknown so an arbitrary number of bins is added. The justification behind this method for colour images is that by increasing the colour range slightly, there would be a greater range of hues of say green thus revealing detail in green regions more clearly. This does indeed occur and does offer a slight improvement on approximate equalization.

Peak Shifting

The reasoning behind this method is similar to that of amplitude scaling. That is, we want to increase the opponent colour difference. The difference between peak shifting and amplitude scaling is that whereas in the latter, the colour of all image pixels is altered by the same factor (m), in peak shifting only the most frequently occurring pixels are changed and then by an amount proportional to their frequency of occurrence. This non-linear method was hoped to better exploit the opponent colour structure.

Peak shifting is based on the fact that since the opponent channels consist predominantly of two colours, a bimodal histogram results. The peaks on this histogram then represent

the two opponent colours. By increasing the distance between these peaks, we increase the opponent colour difference. To do this, we first identify the two peak regions and holding the end points of these regions fixed, we shift the peaks apart. The amount of shift is proportional to the frequency of occurrence or height. Therefore the highest points or the peaks are shifted the furthest. To ensure pixel ordering a smoothed and piecewise monotonic version of the original histogram is used to calculate the shift. This smoothed histogram is obtained using a regularization technique [45].

An example of the use of this technique is shown shown in figure 4.1. Shifting of the two peak regions is clearly visible as well as how pixel ordering has been maintained. The result

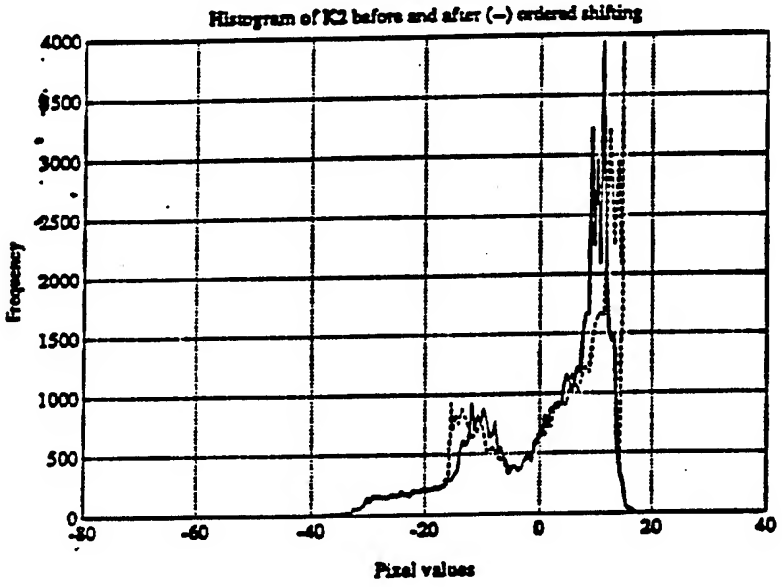


Figure 4.1: Typical opponent channel histogram shown prior to (–) and after (- -) peak shifting

איור 4.1: היסטוגרמה טיפוסית של הערוצים המנוגדים מוצגת לפני (–) ואחרי (- -) הזזת השיאים.

of peak shifting is not significantly different from amplitude scaling.

As neither any of the previous three methods nor amplitude scaling offered satisfactory enhancement, a different approach had to be considered. This is explained in the next two sections.

4.2 Spatial Chromatic Enhancement

The enhancement methods briefly outlined in section 4.1 all ignore the spatial relationships between nearby pixels. Amplitude scaling is an entirely point-wise operation; histogram methods do indirectly affect spatial relations but not in a way that the effect on the visual system is predictable. We know however that spatial relationships are very important to the HVS. We therefore consider using an enhancement technique that emphasizes the most important spatial information. To introduce the technique we first briefly discuss the role in the HVS of retinal lateral inhibition. This discussion reveals what is the most important spatial information and how the HVS obtains it. We then use this knowledge as the basis for our enhancement technique.

4.2.1 The Role of Retinal Lateral Inhibition

We know that spatial information is even more important to the HVS than is chromatic information. Evidence for this includes how much we understand from black and white images or even from simple line drawings. To explain the importance of spatial information, Marr suggested that the Laplacian zero crossings, also known as the primal sketch, of an image are the basis for image understanding by the HVS [46]. The zero crossings of an image are calculated by a Laplacian operator which can also be implemented by centre-surround receptive fields as found in the retina. This therefore provides a rationale for lateral inhibition in the retina. Another rationale for lateral inhibition is provided by the fact that linear predictive coding used by Srinivasan under the hypothesis that the purpose of retinal operations is to reduce redundancy and compress information, predicts the shape of centre-surround receptive fields [24]. Eckert *et al* reconciled these two approaches by showing that the Laplacian and its related zero crossings are, mathematically, a special case of predictive coding. In particular, they showed that the Laplacian is the predictive coder for zero noise images [47].

Thus even though it is unclear whether the importance of contours in image understanding is a byproduct of lateral inhibition used for compression or whether lateral inhibition exists solely to emphasize contours it is clear that contours carry much important information for the HVS. Therefore the basic spatial enhancement method should be one that emphasizes contours and should be performed in a way that mimics the operation of the centre-surround retinal receptive fields.

4.2.2 Enhancement Based on Lateral Inhibition

As we explained in chapter 3, the model separates a digital colour image represented by red, green and blue tristimuli into three other channels, an achromatic luminance channel and two opponent chromatic channels. Although this is done by considering the chromatic content

alone, these channels have a definite spatial structure. Based on the importance of contours to the HVS we propose exploiting this spatial structure for enhancement of the achromatic as well as the chromatic content of the image.

It is well known that in the retina we have luminance opponent centre-surround receptive fields. Thus enhancement based on lateral inhibition on the achromatic channel makes sense. As we discussed in section 2.1, red-green centre-surround receptive fields are also found in the retina and there is evidence for the existence of similar blue-yellow receptive fields, though perhaps in lesser numbers. Thus similar lateral inhibition based processing on the chromatic channels should also enhance the chromatic information in a way compatible with the HVS.

Since LPC predicts centre-surround receptive fields, the best computational approximation of the effect of lateral inhibition in the retina should be given by the use of LPC. Non-causal two dimensional LPC as required by this approach is computationally expensive and unsuitable for real-time applications [48,49]. We know however that the Laplacian is a linear predictor for zero noise images [47]. Thus a good approximation to LPC and thereby retinal lateral inhibition can be obtained if we use the Laplacian to generate zero crossings. This approximation is particularly good for the chromatic channels, K_2 and K_3, because as we showed in section 3.5, these channels are lowpass. Random noise is mainly a high frequency phenomenon so K_2 and K_3 are essentially noiseless. These noise properties are emphasized further in section 4.6.

In terms of image processing the Laplacian is calculated by convolving the image, $K(i,j)$, with a digital approximation of the Laplacian operator. The approximation we use is a standard 3 x 3 mask as follows,

$$L_8 = \frac{1}{8}\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \qquad (4.2)$$

The response of this FIR filter to the ideal step edge shown in figure 4.2 (a) can be seen in figure 4.2 (b). At the position of the edge, a zero crossing response results. If we add this response to the step edge itself, a Mach band type of effect similar to that produced by lateral inhibition by centre-surround receptive fields is obtained. The addition of the edge and its zero crossing are shown in figure 4.2 (c). The Mach band effect can be clearly seen. Based on this discussion, the enhancement we propose to use on each of the model channels is convolution with the Laplacian operator to generate zero crossings and then addition of these zero crossings to the original channels. This enhancement method is illustrated in block diagram form in figure 4.3. The parameter a_i is used to control the amplitude of the generated zero crossings and hence the strength of the perceived Mach bands. The numerical value of this parameter is discussed in section 4.4.

An alternative viewpoint to considering the enhancement to be as a result of increasing

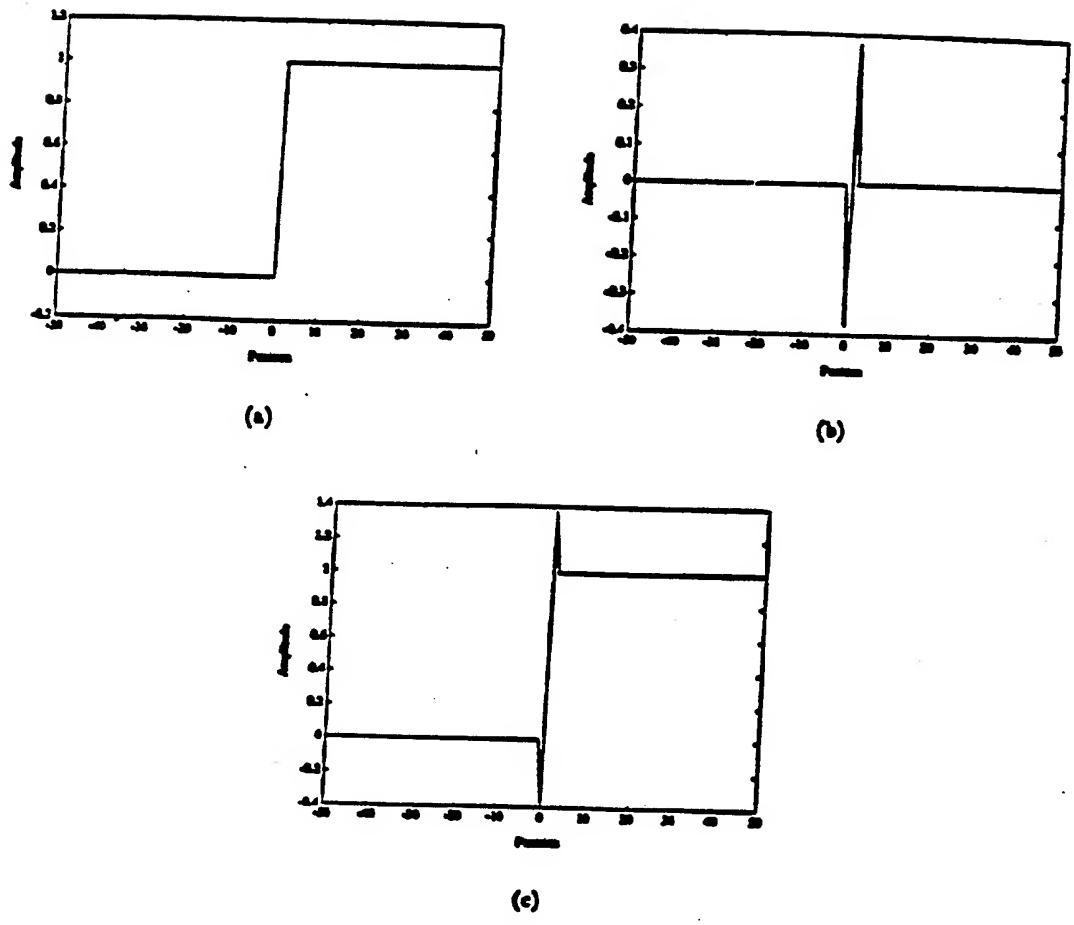


Figure 4.2: The zero-crossing response of the Laplacian FIR filter to the ideal step of (a) is shown in (b). Shown in (c) is the result of adding the zero crossing to the step and the Mach band produced.

איור 4.2: תגובת מסנן הלפלסיאן FIR (Laplacian) למדרגה המוצגת ב (a) מוצגת ב (b). וחיבור תוצרת האפס והמדרגה יוצר גם מך (Mach) המוצג ב (c)

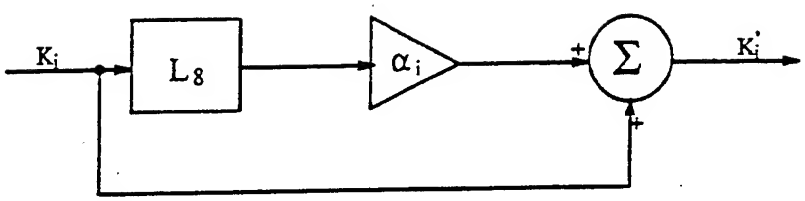


Figure 4.3: Block diagram of the enhancement used on each of the three channels $(i = 1, 2, 3)$
איור 4.3: דיאגרמת מלבנים של תהליך ההדגשה בכל אחד מהערוצים, $(i = 1, 2, 3)$

the sensation of Mach bands may be understood from the following discussion.

A unit impulse $d(i,j)$ may be expressed as,

$$d(i,j) = \frac{8}{9}\left(L_8 + \frac{1}{8}\underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}}_{NA} \right)$$

Any image $f(i,j)$ may then be expressed as,

$$\begin{aligned} f(i,j) &= f(i,j) * d(i,j) \\ &= \frac{8}{9} f(i,j) * L_8 + \frac{1}{9} f(i,j) * NA \\ &= \text{high-pass} + \text{low-pass} \end{aligned} \tag{4.3}$$

where NA is the simple neighborhood averaging mask. Equation 4.3 shows that any image may be divided into a sum of highpass and lowpass components based on 3×3 masks and in particular that the Laplacian component of an image complements the basic spatial lowpass component. By increasing the absolute magnitude of the Laplacian component therefore, we are in effect increasing the amount of high frequency power relative to the low frequency power.

Before we leave this section it is important to discuss why we have chosen to use a small mask. We use a small mask of size less than 9×9 rather than a larger mask since at the current level of technology it is possible to perform very fast convolution operations using pipelined processors only with such small masks [50, 51]. Much research has therefore gone into their use in image processing. The reason we use an even smaller 3×3 mask is that such a small mask is sensitive to very fine detail. Thus even the most insignificant edges are enhanced. This acute sensitivity is essential in the K_2 and K_3 channels. This is because owing to the lowpass nature of K_2 and K_3 edge information in these channels is very reduced. On the other hand, highly sensitive masks tend to increase the image noise. Again because of their lowpass nature, this is not a problem for K_2 and K_3. This may present a problem

for K_1, for as we show in section 4.6, it contains almost all the image noise. To overcome this and still use 3×3 masks, various techniques including lowpass filtering prior to the Laplacian, thresholding the Laplacian generated zero crossings, an image pyramid or even adaptive convolution masks [51], may be used. For standard video pictures of the kind most likely to be encountered, we have not found any of these refinements necessary for visual enhancement of the image.

4.3 An Heuristic Measure of Colour Image Improvement

The results of our enhancement method are shown in the following sections. Visually the effects of enhancement are clear but we would also like to judge these results in a less qualitative and more quantitative manner. To this aim we have investigated a measure of colour image improvement. Ideally this measure should be perceptually based and should be able to rank images according to their quality as perceived by a human viewer. That is, we would like some way of being able to compare an image before and after enhancement and be able to comment on which one is most "visually pleasing". This is an extremely difficult problem whose solution is not even clear for achromatic images. We have therefore had to resort to using a heuristic measure although as will be seen, the measure is based on human perceptual responses.

Since we are enhancing colour images by virtue of their colour, one measure of colour image improvement is an increase in detail revealed by colour. There are two characteristics specific to colour images which we consider important for revealing colour detail. The one is the amount of colour variance or the number of perceptually different colours present in the image. This hypothesis is based on the assumption that an increase in the number of perceptually different colours implies an increase in the number of perceptually different objects in the image and therefore increased detail. The second characteristic is colour contrast by which we mean that the greater the "colour difference" between two objects, the better they are revealed.

Colour Variance

Recently a novel colour quantization method has been attempted [52]. In this method the colour image is first transformed to the almost perceptually uniform *Lab* space. In an ideal uniform space, all colours that lie within a sphere of specified radius are perceived as equivalent. By identifying colour clusters in this space then, the image is requantized so that only perceptually different colours are represented. This method is novel since it does not simply reduce the number of image colours to a number that is possible to display on a graphics device (usually 256) but to that number of colours that are perceptually different. This number of reduced colours is what we are seeking.

Owing to time constraints we have not been able to implement this method however. Results of this method indicate though that an appreciable increase in the number of colours generally also implies an increase in the number of perceptually different colours. A first approximation to this method is thus to use the total number of image colours (N_c), as the first component of the colour measure.

Experience from colour quantization indicates that the effort required to transform from RGB to a uniform space such as Lab or Luv is not justified as better quantization does not result [53]. We have therefore decided to remain in the RGB space when counting the number of colours.

Before counting, we use a technique known as bit slicing to reduce the effect of noise. That is, we requantize the 8 bit R,G and B planes to only 4 bits by slicing off the 4 least significant bits. This implies a reduction of the total number of possible image colours from $(2^8)^3$ to $(2^4)^3$. That is, a variance of 15 within each image plane is considered to be as a result of noise only. This is in effect an initial clustering used to counteract the influence of noise which may be increased as a result of high frequency enhancement. The RGB colour cube is now divided into 4096 sub-cubes. We define the total number of colours as the number of these sub-cubes that are occupied by at least one pixel. The first component of the measure is then the fraction of occupied colour space given by,

$$\rho = \frac{N_c}{4096} \quad \text{where } \frac{1}{4096} \leq \rho \leq 1 \tag{4.4}$$

Average Colour Contrast

The second component of the colour measure is that of the average *colour* contrast. If we have a coloured object on a differently coloured background then the only way to cause the object to stand out from the background, without changing intensity, is to increase the difference in colour between the two. This supposition is based on the Weber-Fechner law which is used to quantify our ability to resolve the same physical stimuli that differ only in intensity. This law states that physical sensation is proportional to log stimulus. That is, the absolute values of two stimuli is not significant for detection but rather their ratio [19,33,12,46].

It is simplest to understand Webers law if we consider the brightness of visual stimuli. Consider a patch of light of brightness $I + \Delta I$ surrounded by a background of intensity I. The difference in brightness, ΔI, is called the just noticeable difference (jnd) when it is the least amount of brightness needed to be added to I to be able to distinguish the object from the background. For a wide range of brightness it has been psychophysically found that

$$\frac{\Delta I}{I} = \text{constant}$$

This result is considered true for a variety of physical sensations and although it has not been verified for colour, it is reasonable to assume that it holds in this case too.

There are two perceptual qualities of colour that may be used to measure colour contrast. These are hue (H) and saturation (S). In most cases where an image is to be enhanced for visual purposes, it is desirable not to distort the values of H since we do not wish to distort the actual colours. Contrast is thus mainly introduced in the saturation image. In addition, to measure colour enhancement alone, we wish to ignore the effects of intensity enhancement. We have therefore chosen to measure colour contrast on the saturation image.

On a complex image that contains more than an object set against a background we have to define a region of interest that can approximate to an object-background scenario. We have found that a 3×3 region centred on each image pixel gives good results. In each of these 3×3 neighborhoods we calculate the Weber fraction as follows:

$$\bar{S}_3(k,l) = \frac{1}{9} \sum_{i=-1}^{1} \sum_{j=-1}^{1} S(i+k, j+l) \tag{4.5}$$

$$\Delta S(k,l) = S(k,l) - \bar{S}_3(k,l) \tag{4.6}$$

$$\bar{W}(k,l) = \frac{\Delta S(k,l)}{S_3(k,l)} \tag{4.7}$$

where

$$\begin{aligned} k &= 2, 3, \ldots, M-1 \\ l &= 2, 3, \ldots, N-1 \\ S(i,j) &= \text{saturation at point (i, j) in each 3 by 3 region} \\ \bar{S}_3(k,l) &= \text{average or background neighborhood saturation, and} \end{aligned}$$

M and N are the number of pixels per row and column respectively, of the image.

Saturation is calculated using a simple cylindrical approximation as follows,

$$S = \sqrt{(R-I)^2 + (B-I)^2}$$

where intensity I is given by,

$$I = 0.3R + 0.59G + 0.11B$$

Finally, the average saturation contrast for the entire image is given as,

$$\bar{W} = \frac{1}{(N-2) \cdot (M-2)} \sum_{k=2}^{N-1} \sum_{l=2}^{M-1} \bar{W}(k,l) \tag{4.8}$$

One should note that there are a number of different ways of defining contrast in images.

These include Michelsons formula for sinusoidal gratings, rms contrast [54], as well as various frequency domain methods [55]. We have chosen Webers approach as it is easy to use, gives good results, corresponds intuitively to the notion of contrast and also seems to be the way that the HVS responds to stimuli differences, at least for brightness.

Total Measure

Both of the components of the measure, ρ and $\bar{W}$, are not absolute measures. That is we cannot comment on the quality of an arbitrary image using them. However as indicators of the *improvement* after colour enhancement we have found that an increase in the values of ρ and $\bar{W}$ does indicate an increase in perceived visual quality. The improvement indicators we therefore use are rather ρ_{rel} and $\bar{W}_{rel}$. These are given by,

$$\rho_{rel} = \frac{\rho_{processed}}{\sigma_{original}} \tag{4.9}$$

and

$$\bar{W}_{rel} = \frac{\bar{W}_{processed}}{\bar{W}_{original}} \tag{4.10}$$

where $\sigma_{original}$ and $\bar{W}_{original}$ are the measures for the original unenhanced image and are given by using equation 4.4 and equation 4.8 respectively on the original image.

Finally, to allocate a single measure of enhancement to the image we need to combine the two components ρ_{rel} and $\bar{W}_{rel}$ such that an increase in either one is reflected in the total. We have found that a simple multiplication of the two components gives adequate results. Therefore the final contrast-density measure we use is,

$$\Gamma = \rho_{rel}\bar{W}_{rel} \tag{4.11}$$

Use of the measure is illustrated in the following sections.

Note that ρ is sensitive to random noise. Thus adding random noise to a picture will increase ρ. We counter this by bit slicing and by restricting the measure to be relative.

4.4 Enhancement of Model Predicted Channels

In this section we test the enhancement procedure described in chapter 3 and illustrated by figure 3.1 on the three channels,K_1, K_2 and K_3, that the transform model predicts. At the same time we discuss guidelines for the selection of values of α_i. At the end of this section, we present in tabular form the values of the colour enhancement measure for each type of enhancement we use.

The discussion of the different aspects of processing will mainly centre upon the "view" image though final results will also be shown on a number of other images. The original of the

"view" image is shown in figure 4.4. It is a standard 512×512 chromatic image comprising 24 bits/pixel. It was obtained using a high quality 3 CCD still video camera with no intentional blurring.

Figure 4.4: Original 512×512 image of the "view". Colour resolution is 24 bits/pixel.
איור 4.4: תמונה מקורית של ה"נוף". הרסולוציה היא 512 × 512 פיקסלים עם 24 סיביות לפיקסל.

4.4.1 Achromatic Processing

Edge enhancement on the intensity component of the image is the standard technique used for colour images. It follows much of our intuitive reasoning of how to enhance edges in colour images as it corresponds with our daily experience of such pictures. For example, simple drawings are usually done by first sketching the outline in pencil and then filling in the patches with colour thereby achieving a colour image with black and white edge enhancement.

To our knowledge, apart from Faugeras [16,17], the only type of edge enhancement that has been performed on colour images is that of the intensity component. Chroma has either been ignored or even blurred by the use of subsampling as a means of image compression [56]. In this section we illustrate the effect of intensity edge enhancement and show why, used on its own, it is deficient.

Intensity edge enhancement is done by enhancing the K_1 channel. When processing the K_1 channel, $K_2' = K_2$ and $K_3' = K_3$ or equivalently, $\alpha_2 = \alpha_3 = 0$.

As we show in section 4.6, the value of α has a direct influence on the increase in noise

power after processing. For this reason, small values of α_1 are desirable. We have found that values in the range of 0.8–1.2 are a good compromise between noise increase and edge enhancement.

The result of enhancement on K_1 only is shown in figure 4.5. A value of $\alpha_1 = 1.0$ was used. Comparing figure 4.4 and figure 4.5, the effect of edge enhancement is clear. The image

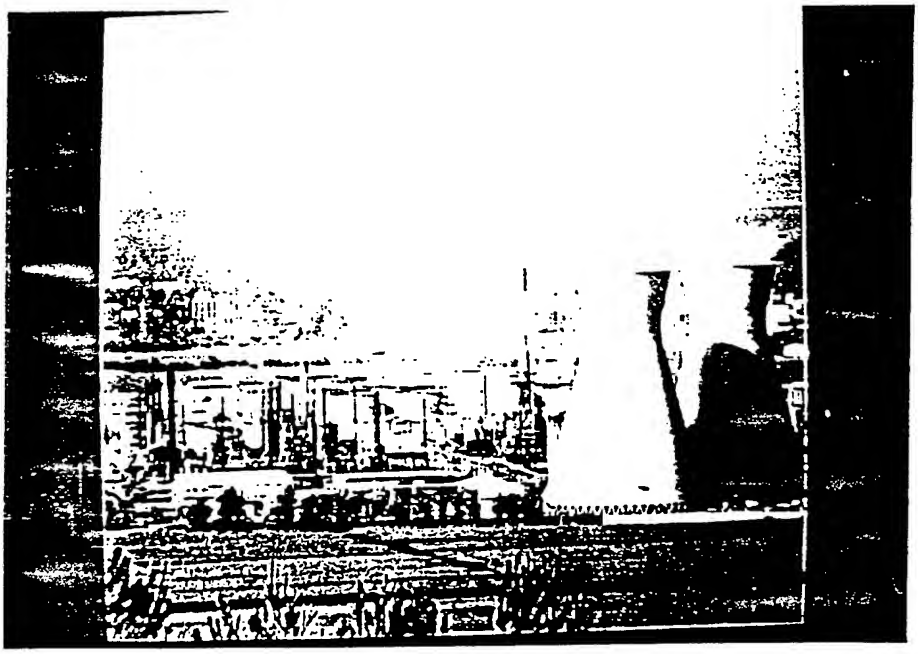


Figure 4.5: Achromatic (K_1) enhancement of "view" image with $\alpha_1 = 1.0$.
איור 4.5: עיבוד אכרומטי (K_1) של ה"נוף" ($\alpha_1 = 1.0$).

appears less blurred with far more visible detail, especially in the background. This increase in background detail gives the image added depth. The cost of this enhancement though is loss of general image colour strength. Colours seem tired and washed out. This reduced colour strength can be explained since it is known that as we increase intensity, saturation at first increases but then rapidly decreases [57]. Another disadvantage of too severe intensity enhancement is that hue changes of as much as 30 nm can occur as a result of the Bezolde-Brucke effect [57]. This means that intensity enhancement alone may not only change the saturation of colours but also their apparent hue. This is undesirable.

4.4.2 Chromatic Processing

Edge enhancement has traditionally not been used on chromatic information for a number of reasons. One reason is that the existence of colour Mach bands in the HVS has not been conclusively shown. Another reason is because of the much reduced sensitivity of the HVS

to higher frequency chromatic contrast as can be seen in figure 2.2. For these reasons, the opposite of edge enhancement, image blurring, has primarily been used on the image colour components to achieve compression [56]. As we saw in the previous section, achromatic enhancement alone is deficient. In this section we show what may be achieved if we are prepared to enhance rather than reduce the high frequency chromatic content of chromatic images.

Convolution with the Laplacian filter produces zero crossings of amplitude proportional to the heights of edges encountered. Since the energy of K_2 is much greater than K_3, edges are more prominent in the K_2 channel. In addition, chromatic contrast sensitivity of the red-green channel is known to be greater than for the yellow-blue channel (see section 2.4). Therefore it is expected that K_2 will be dominant in increasing the perception of high frequency detail. The exact effect of enhancement cannot be predicted and has to be experimentally calculated.

Estimate of α_2 and α_3

Before testing the enhancement of chromatic information, we estimate the values of α_2 and α_3. As already shown, the chromatic channels are spatially lowpass. Therefore they should contain little noise. High frequency enhancement should thus be able to be tolerated to a larger degree than in the K_1 channel. Also, as can be seen from the CSF's shown in figure 2.2, the HVS is far less sensitive to chromatic high frequency information than it is to brightness information. To achieve noticeable high frequency chromatic enhancement therefore, we have to increase the chromatic high frequencies to a greater degree relative to the achromatic ones. This means that we should use larger values of α_2 and α_3 relative to α_1. Without further information we have decided to choose $\alpha_2 = \alpha_3$. To get a feel for the range of α_2 we compared the luminance contrast sensitivity function with the chromatic contrast sensitivity functions of figure 2.2. Where chromatic sensitivity starts to decline, the ratio between these two functions is about 2 and at high frequencies, where the curves are almost parallel, the ratio is about 6. We therefore decided to choose $\alpha_1/\alpha_2 \in [2,6]$.

We are now in a position to perform the chromatic enhancement experiment. Since for the achromatic enhancement of section 4.4.1 we use $\alpha_1 = 1.0$, here we choose $\alpha_2 = \alpha_3 = 3.0$.

Enhancement of K_2 Only:

The result of enhancement on the red-green channel only is shown in figure 4.6. If we compare figure 4.6 to the original of figure 4.4, there is a definite increase in image sharpness. This is particularly evident in areas where there is a colour change such as on the coloured towers and around the foreground bushes. The actual colours are very similar to the original ones. If we now compare this to the result of achromatic enhancement (figure 4.5) we see that edge crispening is not as dramatic but that colour rendering is better. The colours are less washed out.

Figure 4.6: Chromatic enhancement of the K_2 channel of "view" image ($\alpha_2 = 3.0$).
איור 4.6: עיבוד כרומטי של ערוץ (K_2) של ה"נוף" ($\alpha_2 = 3.0$).

Enhancement of K_3 Only:

The result of enhancement on the yellow-blue channel only is shown in figure 4.7. Comparing figure 4.7 to the original (figure 4.4), we notice that the sharpening effect of edge enhancement is very slight. What is dramatic is how at colour edges and around small detail the colour strength is markedly increased. Thus although high frequencies do not increase much, small detail is made more visible by virtue of an increase in strength relative to its surround. This can be clearly seen in the foreground where in the field, rows of plants are suddenly revealed.

Enhancement of K_2 and K_3 Simultaneously:

In figure 4.8 is the result of enhancing both of the chromatic channels simultaneously. Compared to the original, this image exhibits both the sharpness increase contributed by enhancing K_2 as well as the colour detail increase contributed by K_3. What is surprising though is that the sharpness of this image is much sharper than figure 4.6 or figure 4.7 alone. In fact the detail revealed is almost that of the achromatic enhanced image of figure 4.5. This is even more surprising if we remember that together K_2 and K_3 contain less than 1% of the total image energy. Yet as can be seen, their influence on image quality is quite dramatic. Note too that the colour rendering of this image is much better than the achromatic only enhanced image (figure 4.5).

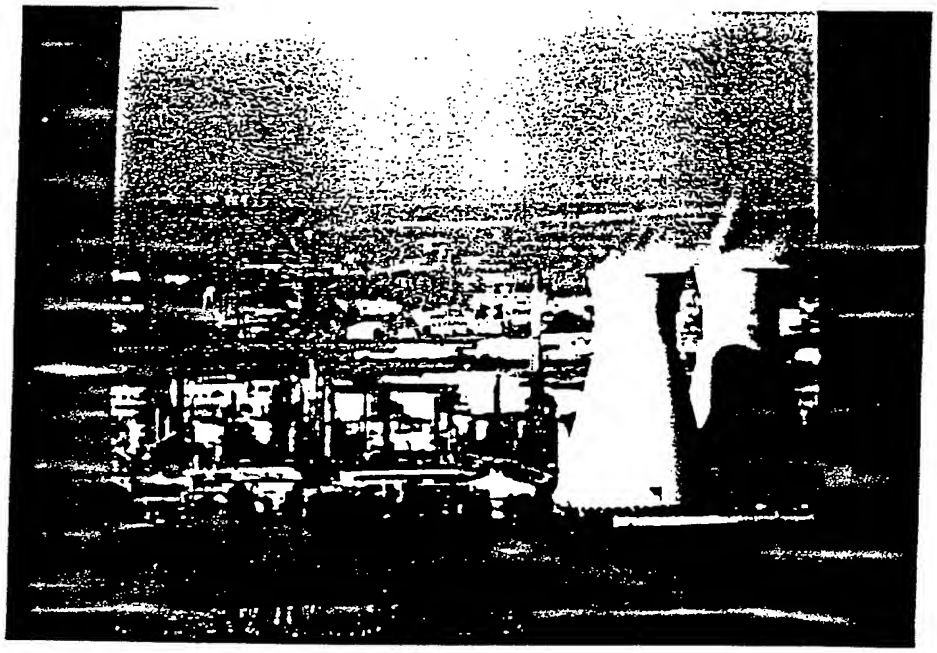


Figure 4.7: Chromatic enhancement of the K_3 channel of "view" image ($\alpha_3 = 3.0$).
איור 4.7: עיבוד כרומטי של ערוץ (K_3) של ה"נוף" ($\alpha_3 = 3.0$).

Enhancement of All Three Channels:

Because of the reduced sensitivity of the HVS to high frequency chromatic contrast, as discussed in section 2.4, very high frequency image enhancement cannot be obtained by processing the chromatic channels alone. Owing to this fact and the promising result of chromatic edge enhancement we are led to believe that the best result would be when we combine achromatic and chromatic enhancement. In this way we would achieve the additional edge crispening that achromatic enhancement achieves and the combination of good colour rendering, small colour detail enhancement as well edge enhancement that chromatic enhancement offers.

To test this we enhanced all three channels simultaneously using $\alpha_1 = 1.0$ and $\alpha_2 = \alpha_3 = 3.0$. The result of this combined enhancement is shown in figure 4.9. It can be seen that this is the sharpest image but still with good colour rendering. This image is also the most visually pleasing from amongst the original and the three prior enhanced images (figures 4.5 to 4.8).

In figure 4.10 and figure 4.11 we show two more examples of the effects of chromatic edge enhancement.

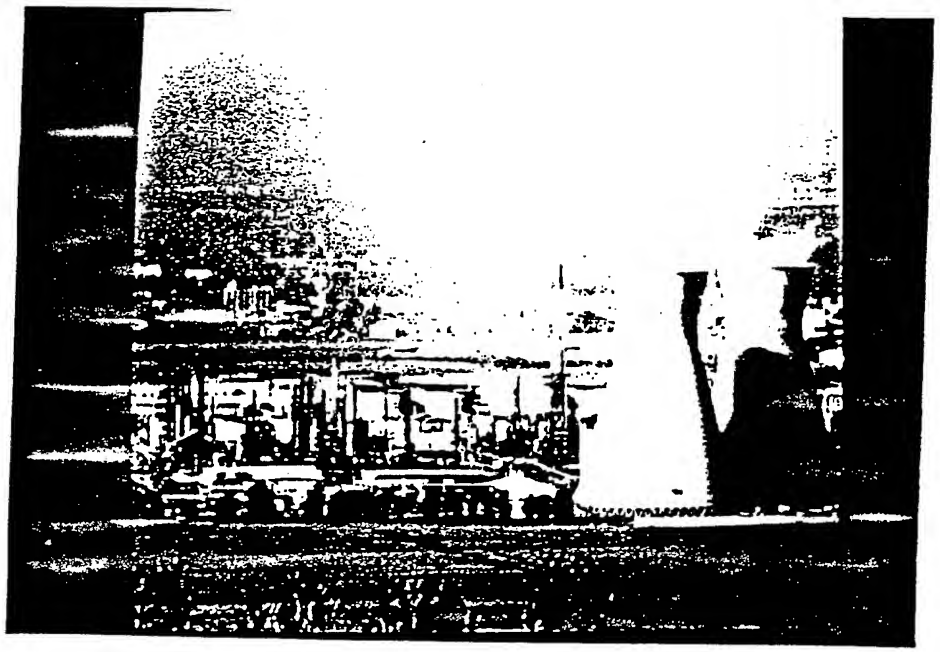


Figure 4.8: Chromatic enhancement of both the K_2 and K_3 channels of "view" image ($\alpha_2 = \alpha_3 = 3.0$).

איור 4.8: עיבוד כרומטי של ערוצי (K_3) ו-(K_2) של ה"נוף" ביחד ($\alpha_2 = \alpha_3 = 3.0$).

Ranking of Results

Here we present for easy reference tabulated results of the quality measures obtained for each experiment. These are shown in table 4.1. The table also contains the colour improvement measures from the results of the methods of section 4.1 that we attempted but that did not produce good results.

From table 4.1 a number of general trends may be noted: Enhancing K_2 and K_3 results in greatly increased colour contrast as well as increased number of colours while enhancing K_1 has little effect. Where no colour detail is improved (amplitude scaling, peak shifting, K_1 enhance), both ρ_{rel} and $\bar{W}_{rel}$ are low while increasing the amount of colour enhancement increases both of these parameters and thus also the total measure Γ. However the actual value of Γ is not a good indicator of visual colour image improvement since it does not reflect for example the marked improvement between the results of the last three entries. Thus although Γ does give an indication of whether colour detail is improved, it is not a reliable measure of the extent of the improvement.

Also included in the table are the results of two control experiments that we performed. In the first experiment we wished to check whether the same results that we obtain by edge enhancing K_2 and K_3 could be obtained by simply edge enhancing the saturation component of the image. This idea stems from the fact that the chromatic components K_2 and K_3 are

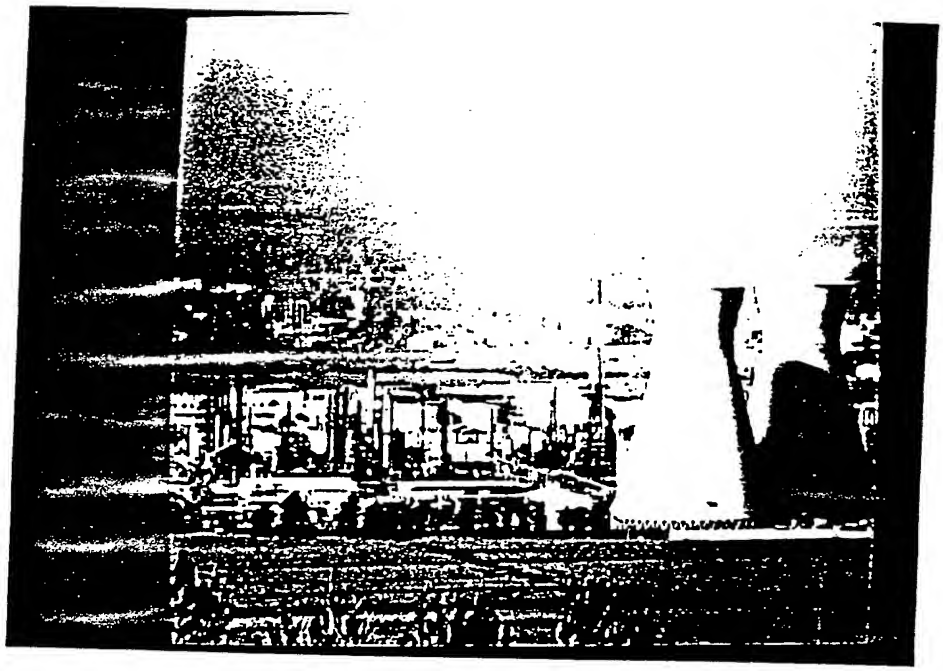


Figure 4.9: Combination of achromatic enhancement on the K_1 channel with $\alpha_1 = 1.0$ as well as chromatic enhancement on both the K_2 and K_3 channels of "view" image using $\alpha_2 = \alpha_3 = 3.0$.

איור 4.9: שילוב של עיבוד אכרומטי על ערוץ K_1 ועיבוד כרומטי על ערוצי הצבע K_2 ו-K_3 של תמונת ה"נוף" $(\alpha_1 = 1.0, \alpha_2 = \alpha_3 = 3.0)$.

directly related to saturation (see section 3.6). To do this, we first transformed from RGB to the hue, intensity and saturation space, HIS. We then edge enhanced S using the Laplacian as we did on the model channels. The result was then transformed back to RGB. Results were poor and did not compare well to the same procedure performed on the model chromatic channels.

In the second experiment we wanted to show that colour detail is not simply enhanced by a mere global saturation increase. To show this, we edge enhanced the K_1 channel and then increased global saturation by a constant multiplicative factor. Again results were poor.

4.5 Enhancement Using a Constant transform

In this section we show how use of the constant transform Φ_c, that we derived in section 3.2.2 for a set of images also results in enhancement that is as good as enhancement using the complete model. It is important to note that the "view" image that we enhance in this section was not used in the derivation of the constant transform but does belong to the same class of natural image scenes. Therefore the enhancement is predicted to be good.

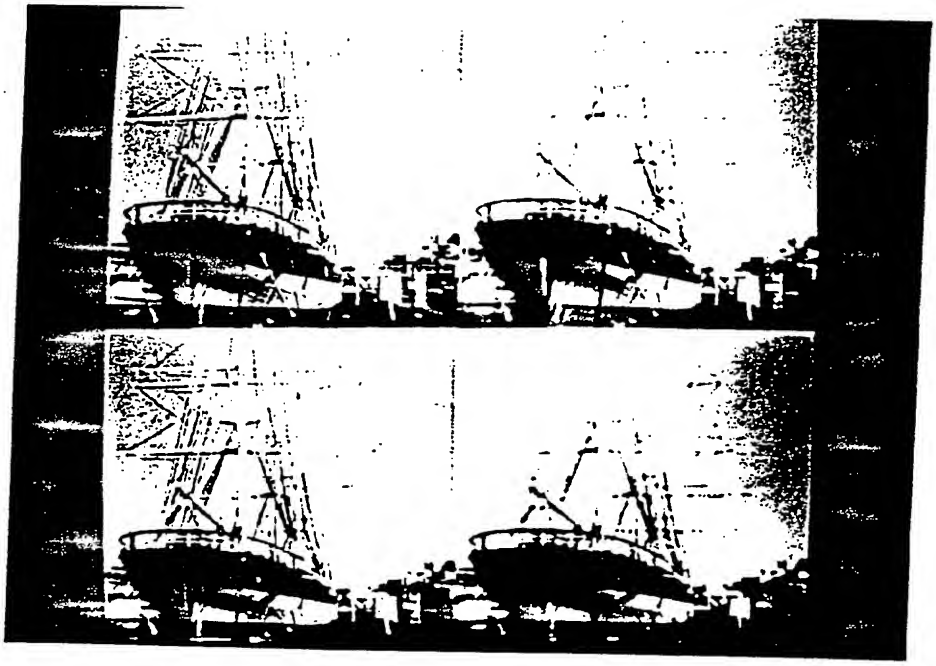


Figure 4.10: Chromatic and achromatic enhancement on "ship" image. Bottom right) original image. Bottom left) Result of K_1 enhancement only. Top right) Result of enhancing K_2 and K_3. Top left) Result of enhancing all three channels simultaneously.

איור 4.10: עיבוד אכרומטי וכרומטי על ה"אונייה". התמונה המקורית מימין למטה, והתוצאה של עיבוד על (K_1) בלבד מלמטה משמאל. מימין למעלה מוצגת התוצאה של עיבוד על (K_2) ו-(K_3) ומשמאל למעלה התוצאה של העיבוד על שלושת הערוצים יחד.

Shown in figure 4.12 is the result of enhancing the "view" image with the same enhancement parameters, $\alpha_1 = 1$, $\alpha_2 = 3$ and $\alpha_3 = 3$ as used in the full transform method of section 4.4.

If we compare figure 4.12 with figure 4.9 we can see that the results are almost identical.

4.6 Enhancement in the Presence of Noise

In this section we show how increasingly large values of α influence the image noise power after processing. We also illustrate how use of our method of transforming the image into an achromatic and two chromatic channels results in almost all the noise being in the achromatic channel and almost no noise in the two chromatic channels. This is due to the lowpass nature of the chromatic channels (see section 3.5). This leads to the conclusion that on noisy image it is preferable to enhance only the chromatic channels. We demonstrate this and show that on very noisy images enhancement on only the chromatic channels produces an image which appears less noisy even when using large values of α.

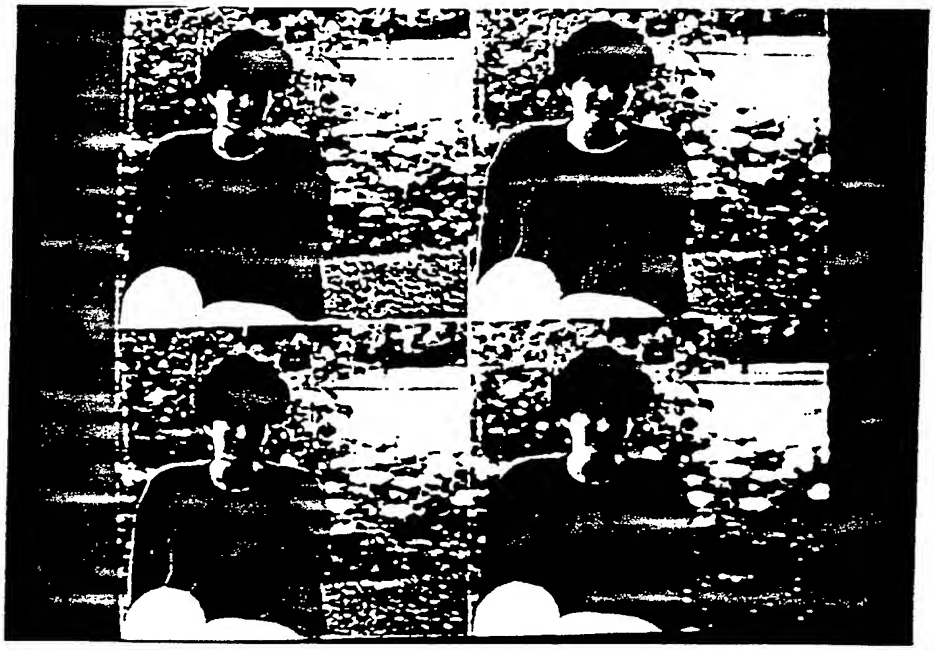


Figure 4.11: Chromatic and achromatic enhancement on "girl" image. Bottom right) original image. Bottom left) Result of K_1 enhancement only. Top right) Result of enhancing K_2 and K_3. Top left) Result of enhancing all three channels simultaneously.

איור 4.11: עיבוד אכרומטי וכרומטי על ה"בחורה". התמונה המקורית מימין למטה, והתוצאה של עיבוד על (K_1) בלבד מלמטה משמאל. מימין למעלה מוצגת התוצאה של עיבוד על (K_2) ו-(K_3) משמאל למעלה התוצאה של העיבוד על שלושת הערוצים יחד.

Before we begin, we state some well known definitions that will help in understanding this section. For a stationary random process $x(n_1, n_2)$, the power spectrum $P_x(\omega_1, \omega_2)$ is defined as,

$$P_x(\omega_1, \omega_2) = \mathcal{F}[R_x(n_1, n_2)]$$

and $R_x(0,0) = \sigma_x^2$ is called the average power of x. For a white noise process,

$$R_x(n_1, n_2; k_1, k_2) = E[x(n_1, n_2)x^*(k_1, k_2)] = \begin{cases} \sigma_x^2 & \text{if } n_1 = k_1, n_2 = k_2 \\ 0 & \text{otherwise} \end{cases}$$

For a stationary white noise process then,

$$\begin{aligned} R_x(n_1, n_2; k_1, k_2) &= E[x(k_1, k_2)x^*(k_1 - n_1, k_2 - n_2)] \\ &= \sigma_x^2 \delta(n_1, n_2) \end{aligned}$$

Operation	ρ_{rel}	W_{rel}	Γ
Peak shift	1.21	1.08	1.30
Amplitude scaling	1.33	1.06	1.40
K_1 enhance	1.14	1.01	1.51
Approx. equaliz.	1.48	1.38	2.04
Smear. approx. equaliz.	1.55	1.40	2.17
Sat. inc. + Inten. edge enh.	1.93	1.15	2.22
Sat. edge enh.	1.85	2.29	4.23
K_3 enhance	1.91	4.15	7.93
K_2 and K_3 enhance	2.95	3.30	9.71
K_2 enhance	2.10	4.69	9.85
K_1, K_2 and K_3 enhance	3.04	3.31	10.06

Table 4.1: Values of the quality indicators in each of the enhancement experiments.
טבלה 4.1: ערכים של מדדי האיכות בכל אחד מניסויי העיבוד.

In other words, the power spectrum of a stationary white noise process is given by

$$P_x(\omega_1, \omega_2) = \sigma_x^2 \ \forall \ (\omega_1, \omega_2) \tag{4.12}$$

Now if we assume that the image noise is white stationary noise of power σ_o^2 then it can be shown that the noise power of the processed image, σ_p^2 is also white stationary noise and is given by,

$$\sigma_p^2 = \sigma_o^2 \sum_i \sum_j h_{i,j}^2 \tag{4.13}$$

where $h_{i,j}$ is the convolution mask used for processing [56, 58].

The enhancement we use, as shown in figure 4.3 can be expressed by,

$$K' = \alpha K * L_8 + K \tag{4.14}$$

Equation 4.14 can be rewritten using the delta mask $d(i,j)$ as,

$$K' = K * (d + \alpha L_8) \tag{4.15}$$

This implies that

$$h_{i,j} = d + \alpha L_8 \tag{4.16}$$

Applying equation 4.13 to expression 4.16 and simplifying results in,

$$\sigma_p^2 = \frac{\sigma_o^2}{8}(9\alpha^2 + 2\alpha + 8) \tag{4.17}$$

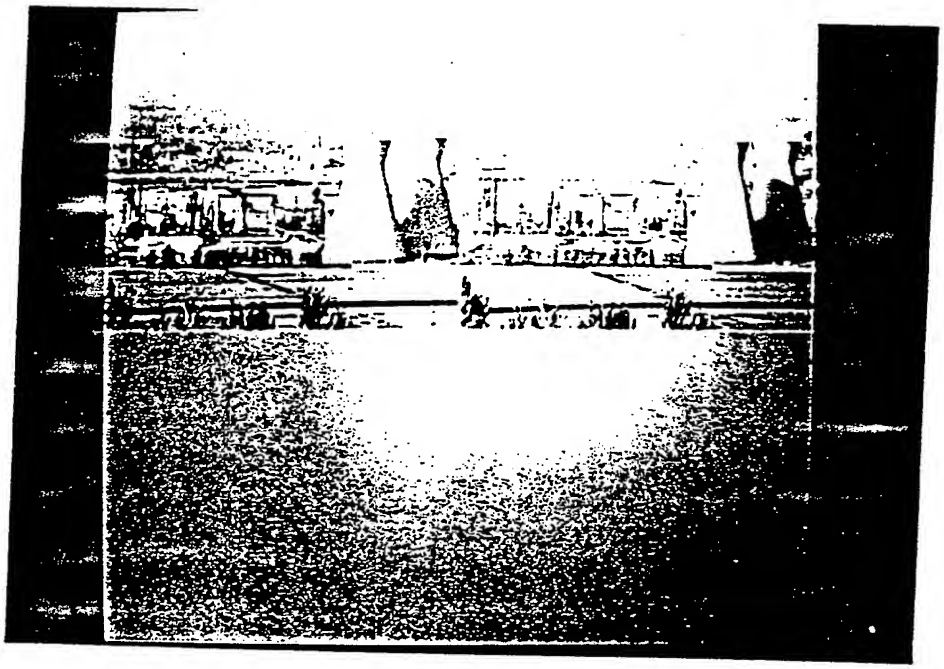


Figure 4.12: Achromatic-chromatic enhancement on "view" image using both the constant transform Φ_c, (top right) and the full computational model (top left).

איור 4.12: עיבוד אכרומטי כרומטי על ה"נוף". התוצאה של עיבוד באמצעות המודל הכללי מוצגת מלמעלה מימין. התוצאה של אותו עיבוד אבל באמצעות ההתמרה הקבועה Φ_c, מוצגת מלמעלה משמאל.

Equation 4.17 shows that an increase in α results in a greater than squared increase in the processed image noise power. For noisy images therefore, α should be kept small.

As an experiment we added to each of the R, G and B planes of the "ship" image stationary white gaussian noise of variance 20. We then transformed this noisy image using the model to obtain the noisy channels K_{1n}, K_{2n} and K_{3n}. We estimated the noise in each channel as

$$N_i = K_{in} - K_i$$

where K_i are the channels obtained when no noise was present. Since the transformation from RGB to the K space is linear, the noise in the K channels will also be stationary white gaussian noise [37]. The resulting noise power will then also be given by equation 4.12.

The resulting noise variance in each of the channels was $\sigma_{k1}^2 = 31$, $\sigma_{k2}^2 = 0.2$ and $\sigma_{k3}^2 = 0.01$. This shows how little noise power is present in the chromatic channels compared to the achromatic one.

We then enhanced only the chromatic channels with $\alpha_2 = \alpha_3 = 3$. If we compare this to the noisy image we can see how due to chromatic enhancement the ship appears to stand out from the background noise and there is no apparent background noise increase. Thus by

seemingly increasing the signal we have increased the signal to noise ratio without decreasing the noise. Apart from the low noise in the chromatic channels, another reason that high frequency enhancement here does not degrade the image is that the HVS is insensitive to very high frequency chromatic contrast.

When we enhanced only the achromatic channel we obtained, as expected, a very much degraded image. Note that we have intentionally not used any noise reduction techniques on the noisy K_1 channel since we only wished to illustrate the benefits of chromatic enhancement and not what may be achieved with further processing.

The results of this experiment are illustrated in figure 4.13. From this we can conclude

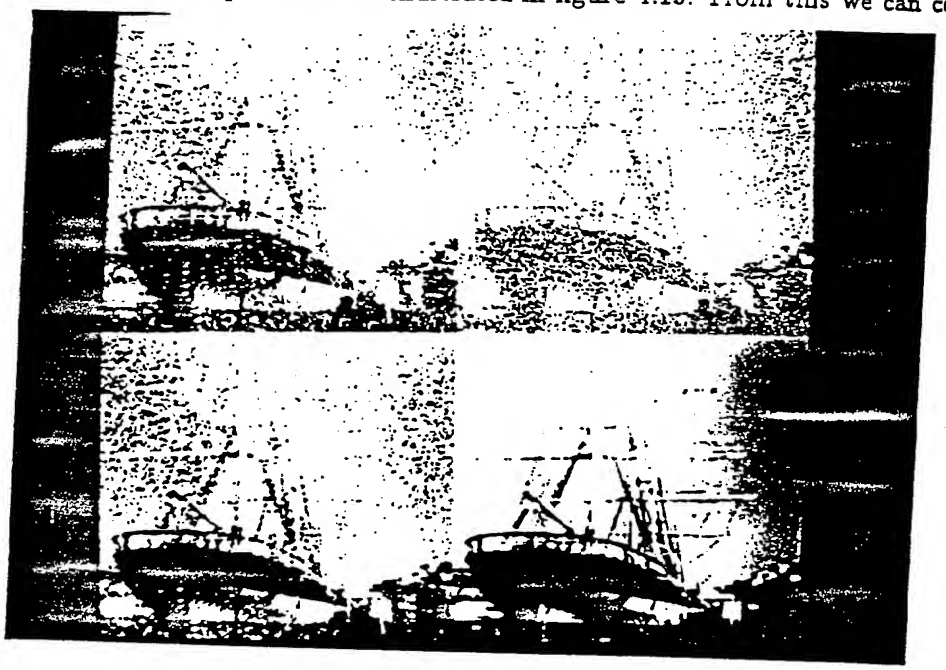


Figure 4.13: Effect of chromatic and achromatic enhancement on noisy ship image. Bottom right) Original image. Bottom left) Unenhanced noisy image. Top left) Result of chromatic enhancement only. Top right) Result of achromatic enhancement only.

איור 4.13: ההשפעה של עיבוד אכרומטי וכרומטי על גירסה רועשת של האוניה. למטה מימין) התמונה המקורית. למטה משמאל) תמונה רועשת. למעלה מימין) תוצאה של עיבוד אכרומטי. למעלה משמאל) תוצאה של עיבוד כרומטי.

that since most images do contain at least some noise, achromatic processing should be used to a limited extent and most enhancement should be performed on the chroma. This is in contrast to the processing techniques currently used. We can also see that noise reduction techniques need only be applied to the achromatic channel and not, as is commonly done, to each of the three planes. This is a great computational saving. Another conclusion is that chromatic edge enhancement can always be used, even on noisy images, without degrading the image. The image may not be enhanced but nor will it be degraded.

Chapter 5

Summary and Conclusions

Summary:

In this report we have developed an opponent colour model of the early HVS for the purpose of colour image processing. We have also investigated some applications of this model to colour image visual enhancement, in particular, colour edge enhancement.

The first stage of the model consists of a linear transformation of a digital image represented by the camera or display tristimuli (RGB) to tristimuli representing the retinal cone responses ($V_l V_m V_s$). By assuming that the difference between the CIE XYZ and modified $\bar{X}\bar{Y}\bar{Z}$ tristimuli is insignificant for image processing, we were able to derive a simple linear transformation T_{RV}. As our results seem to indicate, this assumption was a reasonable one.

The next stage of the model was to transform from the cone responses to the opponent colour space. This transform is not well understood and till now all opponent colour models have been derived by fitting an assumed model to some psychophysical data. We on the other hand have based the transform on new evidence that the HVS obtains its opponent responses in such a way that redundant information is reduced and energy compressed.

In terms of these requirements the Karhunen-Loéve transform is optimal for any ensemble of images. Therefore we use it to derive the opponent transform matrix Φ that reduces *chromatic* correlation. Transforming results in three new planes or channels labelled K_1, K_2 and K_3. The KLT is performed for each image and even though it is simpler to perform than spatial decorrelation, the operation still takes about 1.5 minutes on a Sun-4. This generalized model is thus not suitable for real time use.

We also showed however that it is possible to derive a constant transform Φ_c that performs as well as the specific transform for any set of similar images. Since we wished to show at the same time that our model approximates to the early HVS we decided to derive Φ_c for the ensemble of natural daylight images. We did this because we firstly assume that the HVS opponent transform is constant. This is because it is performed in real time in the retina. Secondly, we assume that as man has adapted optimally to his environment, the transform

for the natural daylight group of images should be similar to the HVS opponent transform. A Karhunen-Loéve analysis of such a set of images yielded Φ_c. We showed that Φ_c is as good as the specific image transform for a human viewer by showing that the same enhancement using both transforms yields results whose difference is imperceptible. The implication of this is that the generalized model can be reduced to a much simplified model involving only two linear transforms represented by two 3×3 matrices. This model can be realized in real time.

To show the similarity of our model to the HVS, we transformed the cone response curves using Φ_c. This yielded three new spectral response curves. The response curve for the K_1 channel was very similar to the relative luminous efficiency curve that has been psychophysically determined. We thus call K_1 the achromatic or Black-White channel. The spectral response of the K_2 channel has opponent peaks in the red-green area of the spectrum and is thus called the Red-Green opponent channel. For similar reasons, K_3 is called the Blue-Yellow opponent channel. These opponent spectral response curves are also similar to those psychophysically predicted. This illustrates the similarity of our model to the HVS.

Investigating some of the properties of the model channels showed that K_1 is spatially highpass and contains almost all the image energy. K_2 and K_3 are spatially lowpass and contain very little energy.

As a side issue, we noted that when plotted on a log-log scale the achromatic channel for natural images exhibits approximately $1/f$ behaviour as has previously been found. It is intersting to note though that the chrominance channels behave differently to the achromatic channel. Their behaviour is approximately of the form $1/2f$.

We then considered enhancing in the opponent K space and then transforming back to RGB space to view the results. On achromatic images, edge enhancement has been used for a long time with good results. This is due to the existence of Black-White opponent centre-surround retinal receptive fields implying luminance lateral inhibition and hence the importance of edge information. Until now this same method of enhancing has only been used on the intensity component of chromatic images. In fact colour has usually been blurred especially for coding purposes. Because of recent evidence of opponent colour centre-surround receptive fields we investigated performing colour edge enhancement too. This is done by simulating lateral inhibition by determining the Laplacian zero crossings of the relevant channel and adding these to that same channel. This has the effect of increasing perception of Mach bands. The Laplacian zero crossings are calculated using a small (3×3) digital Laplacian mask. Therefore the enhancement procedure is also suitable for real time operation.

The amount of enhancement is controlled by a parameter α and we have found that for achromatic enhancement, $\alpha = 1$ while for chromatic enhancement $\alpha = 3$ are reasonable values.

Achromatic enhancement performed on the K_1 channel produces as expected a much sharper image. At edges though, because of the increased intensity, colours seem faded thus making the whole image appear washed out. Enhancement performed on K_2 alone produces

an image with better colours that is slightly sharper than the original. Enhancement on K_3 produces no sharpness increase but small coloured detail is made to stand out by increasing its apparent colour strength. The most surprising result is that enhancing both chromatic channels simultaneously produces an image which is almost as sharp as that obtained when achromatic enhancement is used but with the additional property of richer colours. This is even more surprising when we recall that the chromatic channels contain less than 1 % of the total image energy. The best result is obtained when we combine achromatic and chromatic enhancement as then we benefit from the increased high frequency enhancement that because of the nature of the HVS, chromatic enhancement alone cannot provide.

The same enhancement performed in spaces other than the HVS K space did not produce comparable results and even degraded the image. This seems to indicate that without a quantitative definition of subjective image quality, enhancement should be based on a visual model and that our visual model is a good one for image processing purposes.

We also investigated the effect of enhancement on noisy images. We were concerned with this since edge enhancement generally increases high frequency noise. By adding white gaussian uncorrelated noise to each of the RGB planes and then transforming we estimated the noise in each of the K channels. This showed that K_1 contains nearly all the noise and K_2 and K_3 negligibly little. This is understandable because of the highpass and lowpass nature of the different channels. Enhancement of the chromatic channels alone thus does not increase the noise much but does seem to improve the image therefore improving somehow subjective signal to noise ratio. An implication of this is that chromatic edge enhancement can always be used without fear of degrading the image.

Finally, we investigated a quantitative measure of colour image enhancement. Presently the only way of evaluating results has been to use subjective responses. We wanted to use a quantitative measure of post processing colour image improvement. Due to the highly complex nature of human colour vision we attempted to simplify matters greatly and based this measure on two features that we considered to be important for colour detail improvement. These were the number of perceptually different colours and average colour contrast. An increase in both of these should indicate an increase in visible colour detail. Experimentally we found that this is true but that the absolute values of these parameters are not a reliable way of comparing one method against another. All they can tell you is whether there was some improvement or not but not the level of improvement. The measure is therefore not reliable.

Conclusions:

- The correspondence between the model response curves in the case of natural images and those psychophysically predicted demonstrate the similarity of the model opponent process and that found in the HVS.

- Chromatic edge enhancement produces results not possible by achromatic edge enhancement alone. We therefore feel that when dealing with colour images it is worth exploiting the additional special qualities of such images and to perform chromatic edge enhancement as well as the standard intensity enhancement procedures.
- Since it is possible to simplify the opponent transform from a Karhunen-Loéve transform per image to a constant transform for a group of images, the model can be realized in real time.
- Assuming that a similar redundancy reduction process is performed in colour blind or partially colour blind people (protanopes or deuteranopes), the model can be tuned to enhance images for the benefit of such people. This may be done by simply redefining the cone response curves such that they match those of the colour deficient person.
- The fact that most image noise is contained in the achromatic channel makes chromatic edge enhancement possible without fear of degrading the image. Thus chromatic edge enhancement can be incorporated as a permanent feature of image capture and display systems.
- The simplicity of the enhancement procedure (3×3 mask convolution) means that it may also be used in real time.
- The measure of image quality is unsatisfactory and cannot be used as a reliable estimate of post processing image quality.
- From the amplitude spectra it seems that natural images can be typified by both the behaviour of the achromatic component ($1/f$) and the chromatic components ($1/2f$).

Further work:

There are a few areas which should be investigated further:

- We know that the HVS not only reduces chromatic correlation but also spatial correlation. This effect should be incorporated into the opponent transform perhaps by performing a neighborhood Karhunen-Loéve transform rather than by performing it pointwise. Non-linearities of the HVS could also be incorporated.
- We have only looked at one application of the model in colour image processing. An immediate further application is to the problem of quantization. Because of the way energy is spread amongst the three channels, simple and efficient quantization may be possible by allocating the most bits to the luminance channel, less to the red-green channel and the least to the blue-yellow channel. Other possible applications are coding and noise reduction. Applications of the model to the problem of colour constancy should also be investigated.

- The measure of enhancement is far from comprehensive and requires much more work. Without changing the measure, two immediate extensions would be to use a larger neighborhood for estimating background saturation and to weight ρ by the number of pixels of each colour so that less occupied cubes in colour space would be less significant.

References

[1] G. Buchsbaum, "A spatial processor model for object colour perception," *Journal of the Franklin Institute*, vol. 310, pp. 1–26, July 1980.

[2] R. Gershon and A. D. Jepson, "The computation of color constant descriptors in chromatic images," *COLOR research and application*, vol. 14, pp. 325–334, Dec. 1989.

[3] L. T. Maloney and B. A. Wandell, "Color constancy: A method for recovering surface spectral reflectance," *Journal of the Optical Society of America A*, vol. 3, pp. 29–33, Jan. 1986.

[4] E. H. Land and J. J. McCann, "Lightness and retinex theory," *Journal of the Optical Society of America A*, vol. 61, pp. 1–11, Jan. 1971.

[5] E. H. Land, "The retinex theory of color vision," *Sci. Amer.*, vol. 237, pp. 108–128, 1977.

[6] E. H. Land, "An alternative technique for the computation of the designator in the retinex theory of color vision," *Proc. Natl. Acad. Sci. USA*, vol. 83, pp. 3078–3080, May 1986.

[7] A. Moore, J. Allman, and R. M. Goodman, "A real-time neural system for color constancy," *IEEE Trans. on Neural Networks*, vol. 2, pp. 237–247, Mar. 1991.

[8] B. R. Hunt and O. Kubler, "Karhunen-loeve multispectral image restoration, part 1: Theory," *IEEE Trans. Accoustics, Speech, and Signal Processing*, vol. ASSP-32, pp. 592–599, June 1984.

[9] R. C. Gonzalez and P. A. Wintz, *Digital Image Processing.* Addison-Wesley, second ed., 1987.

[10] W. K. Pratt, *Digital Image Processing.* John Wiley and Sons, 1978.

[11] A. Rosenfeld and A. C. Kak, *Digital Picture Processing.* Academic Press, second ed., 1982.

[12] J. S. Lim, *Two-Dimensional Signal and Image Processing.* Prentice-Hall, 1990.

[13] A. F. Lehar and R. J. Stevens, "High-speed manipulation of the color chromaticity of digital images," *IEEE Trans. on Computer Graphics and Animation*, pp. 34–39, Feb. 1984.

[14] R. N. Strickland, C. Kim, and W. F. Mcdonnell, "Digital color image enhancement based on the saturation component," *Optical Engineering*, vol. 26, pp. 609–616, July 1987.

[15] I. Bockstein, "Color equalization method and its application to color image processing," *Journal of the Optical Society of America A*, vol. 3, pp. 735–737, May 1986.

[16] O. D. Faugeras, *Digital Color Image Processing and Psychophysics within the Framework of a Human Visual Model.* PhD dissertation, University of Utah, Department of Computer Science, June 1976.

[17] O. D. Faugeras, "Digital color image processing within the framework of a human visual model," *IEEE Trans. Accoustics, Speech, and Signal Processing*, vol. ASSP-27, pp. 380–393, Aug. 1979.

[18] G. Wysecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitive Data and Formulae.* John Wiley and Sons, second ed., 1982.

[19] R. M. Boynton, *Human Color Vision.* Holt, Rinehart and Winston, 1979.

[20] G. Buchsbaum and A. Gottschalk, "Chromaticity coordinates of frequency-limited functions," *Journal of the Optical Society of America A*, vol. 1, pp. 885–887, Aug. 1984.

[21] E. Hering, *Outlines of a Theory of the Light Sense.* Harvard University Press, 1964. Translated by Hurvich and Jameson.

[22] R. Gershon, "Aspects of perception and computation in color vision," *Computer Vision, Graphics, and Image Processing*, vol. 32, pp. 244–277, 1985.

[23] S. L. Guth, "Model for color vision and light adaption," *Journal of the Optical Society of America A*, vol. 8, pp. 976–993, June 1991.

[24] M. V. Srinivasan, S. B. Laughlin, and A. Dubs, "Predictive coding: A fresh view of inhibition in the retina," *Proceedings of the Royal Society of London B*, vol. 216, pp. 427–939, 1982.

[25] Y. Tsukamoto, R. G. Smith, and P. Stirling, "Collective coding of correlated cone signals in the retinal ganglion cell," *Proc. Natl. Acad. Sci. USA*, vol. 87, pp. 1860–1864, Mar. 1990.

[26] G. Buchsbaum and A. Gottschalk, "Trichromacy, opponent colours coding and optimum colour information transmission in the retina," *Proceedings of the Royal Society of London B*, vol. 220, pp. 89–113, 1983.

[27] J. B. Derrico and G. Buchsbaum, "A computational model of spatichromatic image coding in early vision," *Journal of Visual Communication and Image Representation*, vol. 2, pp. 31–38, Mar. 1991.

[28] A. M. Rohaly and G. Buchsbaum, "Inference of global spatiochromatic mechanisms from contrast sensitivity functions," *Journal of the Optical Society of America A*, vol. 5, pp. 572–576, Aug. 1988.

[29] A. N. Netravali and B. G. Haskell, *Digital Pictures.* 233 Spring Street, New York, N.Y. 10013: Plenum Press, 1988.

[30] G. Buchsbaum, "Color signal coding: Color vision and color television," *COLOR research and application*, vol. 12, pp. 266–269, Oct. 1987.

[31] J. J. Vos, "Colorimetric and photometric properties of a 2° fundamental observer," *COLOR research and application*, vol. 3, pp. 125–128, 1978.

[32] A. K. Jain, "Advances in mathematical models for image processing," *Proceedings of the IEEE*, vol. 69, pp. 502–528, May 1981.

[33] A. K. Jain, *Fundamentals of Digital Image Processing.* Prentice Hall, Inc, 1989.

[34] W. K. Pratt, "Spatial transform coding of color images," *IEEE Trans. on Communication Technology*, vol. COM-19, pp. 980–992, Dec. 1971.

[35] F. M. Wahl, *Digital Image Signal Processing.* 685 Canton Street, Norwood, MA 02062, England: Artech House Inc., 1987.

[36] P. A. Wintz, "Transform picture coding," *Proceedings of the IEEE*, vol. 60, pp. 809–820, July 1972.

[37] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, second ed., 1984.

[38] J. A. Saghri and A. G. Tescher, "Near-lossless bandwidth compression for radiometric data," *SPIE-The International Society for Optical Engineering*, vol. 30, pp. 934–939, July 1991.

[39] T. G. Stockham, "Image processing in the context of a visual model," *Proceedings of the IEEE*, vol. 60, pp. 828–842, July 1972.

[40] C. F. Borges, "Trichromatic approximation method for surface illumination," *Journal of the Optical Society of America A*, vol. 8, pp. 1319–1323, Aug. 1991.

[41] W. Frei and B. Baxter, "Rate-distortion coding simulation for color images," *IEEE Trans. on Communication Technology*, vol. COM-25, pp. 1385–1392, Nov. 1977.

[42] D. M. Schneider, M. Steeg, and F. H. Young, *Linear Algebra, A Concrete Introduction.* Macmillan Publishing Co., Inc., 1982.

[43] D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing.* Prentice-Hall, 1984.

[44] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *Journal of the Optical Society of America A*, vol. 4, pp. 2379–2394, Dec. 1987.

[45] P. Srimanta and P. Bhattacharyya, "Multipeak histogram analysis in region splitting: a regulisation problem," *IEE Proceedings-E*, vol. 138, pp. 285–288, July 1991.

[46] D. Marr, *Vision.* Freeman, 1982.

[47] M. P. Eckert, J. B. Derrico, and G. Buchsbaum, "The laplacian zero-crossings of images are a special case of predictive coding in the retina," *Invest. Opthalmol. Visual Sci.*, vol. 30(Suppl), no. 24, 1989.

[48] J. Makhoul, "Linear prediction: A tutorial review," *Proceedings of the IEEE*, vol. 63, pp. 561–580, Apr. 1975.

[49] P. A. Maragos, R. W. Schafer, and R. M. Mersereau, "Two-dimensional linear prediction and its application to adaptive predictive coding of images," *IEEE Trans. Accoustics, Speech, and Signal Processing*, vol. ASSP-32, pp. 1213–1229, Dec. 1984.

[50] J. E. Hall and J. D. Awtrey, "Real-time image enhancement using 3 by 3 pixel neighbourhood operator functions," *Optical Engineering*, vol. 19, pp. 421–424, May 1980.

[51] R. N. Strickland and M. Y. Aly, "Image sharpness enhancement using adaptive 3 by 3 convolution masks," *Optical Engineering*, vol. 24, pp. 683–686, July 1985.

[52] A. Tremeau, B. Laget, and M. Calonier, "Automatic color quantization through a perceptual approach," *Journal of Visual Communication and Image Representation*, vol. 2, pp. 2032–2040, Oct. 1992. to be published.

[53] N. Goldberg, "Colour image quantization for high resolution graphics display," *Image and Vision Computing*, vol. 9, pp. 303–312, Oct. 1991.

[54] M. W. Cannon, "Perceived contrast in the fovea and periphery," *Journal of the Optical Society of America A*, vol. COM-19, pp. 1760–1768, Oct. 1985.

[55] E. Peli, "Contrast in complex images," *Journal of the Optical Society of America A*, pp. 2032–2040, Oct. 1990.

[56] V. R. Algazi, "Fir anisotropic filters for image enhancement," in *ICASSP-86: Proceedings*, 1986.

[57] H. J. Durret, ed., *Color and the Computer.* Academic Press, 1991.

[58] V. R. Algazi, G. E. Ford, and E. Hildum, "Digital representation and storage of high quality color images by anisotropic enhancement and subsampling," in *ICASSP-89: Proceedings*, 1989.

Appendix A

Colourimetry

As we have seen in chapter 2 the initial cone responses are only the first stage in determining the colour of an object. Notwithstanding this and the fact that colour processing in the HVS is poorly understood, there is a need to deal with colour in fields as diverse as television and paint manufacture. Therefore to standardize the naming of colours the CIE (Commission de L'Eclairage) adapted a standard based on the assumption of trichromacy.

Trichromacy has been taken to imply two basic assumptions also known as Grassmans Laws,

- Any colour can be matched by a vectorially additive combination of three primaries.
- Colour processing is linear, indicating the properties of proportionality and additivity.

The CIE standard transforms radiant energy into a three dimensional colour space. The equations governing this transformation are:

$$L = \int_{\lambda} I(\lambda) S_l(\lambda) d\lambda \quad \text{(A.1)}$$

$$M = \int_{\lambda} I(\lambda) S_m(\lambda) d\lambda \quad \text{(A.2)}$$

$$S = \int_{\lambda} I(\lambda) S_s(\lambda) d\lambda \quad \text{(A.3)}$$

where the scalars L,M,S are the tristimulus values and S_l, S_m and S_s are three hypothesized colour filters which play a role similar to that of the colour absorption curves of the cone pigments.

To determine S_l, S_m and S_s, colour matching experiments using three linearly independent monochromatic primaries are used. The curves thus depend on the choice of primaries. In order to standardize and avoid negative values response values, the CIE defined the well known XYZ system which is based on a set of unreal primaries. As measurement techniques improved these 1931 CIE XYZ values were modified by Judd in 1964 to obtain the $\bar{X}\bar{Y}\bar{Z}$ values.

A digital image is usually represented by three planes containing the tristimulus values related to the set of primaries used to display the image. These are usually Red, Green and Blue.

Appendix B

Scene Independent Transform

We wish to find an opponent colour transform that can be used for pictures of a similar type. To do this we propose calculating a Karhunen-Loéve transform not on a single image but on a set (Y) of images. To calculate the Karhunen-Loéve transform we first have to estimate the covariance matrix of Y. It is not feasible to simply make one super-image consisting of all the images in the set and then estimate its covariance matrix owing to the immense amount of data this will entail. In what follows we derive an estimate of the covariance of the set of images, C_y.

In order to simplify the mathematical notation used further on we will now consider each L by W image plane as a one dimensional 1 by $L \cdot W$ vector. Multispectral images consist of M planes while standard colour images have $M = 3$. For each image we may write

$$X_j = \begin{bmatrix} x_{j,1} \\ x_{j,2} \\ \vdots \\ x_{j,M} \end{bmatrix}$$

where $x_{j,m}$ is the intensity of plane m at location j in the 1 by $L \cdot W$ vector. The full multispectral image may then be written as

$$\mathbf{X} = \begin{bmatrix} X_1 & X_2 & \cdots & X_j & \cdots & X_{LW} \end{bmatrix}$$

Note that it is unimportant whether we form each vector as a row or column stack. With the new notation in mind we may rewrite equation 3.5 as:

$$\mathbf{C_x} \approx \frac{1}{L \cdot W} \sum_{j=1}^{LW} (X_j - \mathbf{m_x})(X_j - \mathbf{m_x})^t \qquad \text{(B.1)}$$

where we similarly rewrite equation 3.6 as

$$\mathbf{m}_x \approx \frac{1}{L \cdot W} \sum_{j=1}^{LW} X_j \tag{B.2}$$

If we expand equation B.1 we obtain:

$$\mathbf{C_x} \approx \frac{1}{L \cdot W} \sum_{j=1}^{LW} \left(X_j X_j^t - X_j \mathbf{m}_x^t - \mathbf{m}_x X j^t + \mathbf{m}_x \mathbf{m}_x^t \right)$$

and noting that

$$-\left(\frac{1}{L \cdot W} \sum_{j=1}^{LW} X_j \right) \mathbf{m}_x^t - \mathbf{m}_x \frac{1}{L \cdot W} \sum_{j=1}^{LW} X_j^t \;=\; -\mathbf{m}_x \mathbf{m}_x^t - \mathbf{m}_x \mathbf{m}_x^t$$

and that

$$\frac{1}{LW} \sum_{j=1}^{LW} \mathbf{m}_x \mathbf{m}_x^t \;=\; \mathbf{m}_x \mathbf{m}_x^t$$

equation B.1 then simplifies to:

$$\mathbf{C_x} \approx \left(\frac{1}{L \cdot W} \sum_{j=1}^{LW} X_j X_j^t \right) - \mathbf{m}_x \mathbf{m}_x^t \tag{B.3}$$

Now consider the set of data to consist of N images each of size L_n by W_n where $\mathrm{n} = 1 \ldots N$. In keeping with the vector notation, each element of Y can be written as:

$$Y_{n,j} = \begin{bmatrix} y_{n,j,1} \\ y_{n,j,2} \\ \vdots \\ y_{n,j,M} \end{bmatrix}$$

In order to simplify the development we order Y such that all the $Y_{1,j}$s are first then the $Y_{2,j}$s and so on. That is, all the pixel vectors of each image are grouped together. This results in

$$\mathbf{Y} = \left[\; Y_{1,1} \;\; \cdots \;\; Y_{1,L_1 W_1} \;\; Y_{2,1} \;\; \cdots \;\; Y_{2,L_2 W_2} \;\; \cdots \;\; Y_{N,1} \;\; \cdots \;\; Y_{N,L_N W_N} \; \right]$$

The total number of vectors in Y is

$$d = \sum_{n=1}^{N} L_n W_n \tag{B.4}$$

On occasion it will be simpler to index the elements of $\mathbf{Y}$ by a single index j where $j = 1 \ldots d$. Thus it is to be understood that $\mathbf{Y}_j$ is the j'th vector of $\mathbf{Y}$.

If we now apply equation B.3 to set $\mathbf{Y}$ we get

$$\begin{aligned} \mathbf{C}_y &= \left(\frac{1}{d}\sum_{j=1}^{d} Y_j Y_j^t\right) - \mathbf{m}_y \mathbf{m}_y{}^t \\ &= \frac{1}{d}\left(\sum_{n=1}^{N}\sum_{i=1}^{L_n W_n} Y_{n,i} Y_{n,i}^t\right) - \mathbf{m}_y \mathbf{m}_y{}^t \end{aligned} \tag{B.5}$$

but also from equation B.3 we see that

$$\sum_{j=1}^{L_n W_n} Y_{n,j} Y_{n,j}^t = L_n W_n \left(\mathbf{C}_n + \mathbf{m}_n \mathbf{m}_n{}^t\right) \tag{B.6}$$

In addition, if we apply equation B.2 to the set of data we get

$$\begin{aligned} \mathbf{m}_y &= \frac{1}{d}\sum_{j=1}^{d} Y_j \\ &= \frac{1}{d}\sum_{n=1}^{N}\sum_{i=1}^{L_n W_n} Y_{n,i} \\ &= \frac{1}{N}\sum_{n=1}^{N}\left(\frac{1}{L_n W_n}\sum_{i=1}^{L_n W_n} Y_{n,i}\right) \\ &= \frac{1}{N}\sum_{n=1}^{N} \mathbf{m}_n \end{aligned} \tag{B.7}$$

where m_n is the mean of each image in the set. That is, the mean of the set of images is the mean of the means of each image.

Now substituting equations B.6 and B.7 into equation B.5

$$\mathbf{C}_y = \frac{1}{d}\sum_{n=1}^{N} L_n W_n \left(\mathbf{C}_n + \mathbf{m}_n \mathbf{m}_n{}^t\right) - \frac{1}{N^2}\left(\sum_{n=1}^{N} \mathbf{m}_n\right)\left(\sum_{n=1}^{N} \mathbf{m}_n{}^t\right) \tag{B.8}$$

If we restrict ourselves to images of equal size that is, $L_1 W_1 = L_2 W_2 = \ldots L_N W_N = LW$ then from B.4,

$$d = NLW$$

and equation B.5 simplifies to

$$\mathbf{C}_y = \frac{1}{N}\left[\sum_{n=1}^{N}\left(\mathbf{C}_n + \mathbf{m}_n \mathbf{m}_n{}^t\right) - \frac{1}{N}\left(\sum_{n=1}^{N} \mathbf{m}_n\right)\left(\sum_{n=1}^{N} \mathbf{m}_n{}^t\right)\right] \tag{B.9}$$

Chromatic Enhancement Processor

Hardware Description

The CE processor is described in Figure 1. It comprises a color converter 12, a decorrelating transform calculator 80, a decorrelating transformer 82, enhancement processor 40, inverse decorrelating transformer 50, and inverse color converter 60.

COLOR CONVERTER 12

The color converter 12 converts the acquired color signal, typically RGB or YC_rC_b, to a representation similar to that of biological retinal receptors. One possible embodiment of this converter comprises a 3×3 matrix multiplier, converting the three color components (e.g. R, G, B) at each image point to three new components V_l, V_m, V_s, according to a linear transform. The rationale for such embodiment, as well as the derivation of the transform values, are detailed in Stuart Wolf's thesis, "Development of an Adaptive Opponent Color Model for Color Image Processing," presented to the Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa, Israel, in June 1992 [W92]. The transform values are given below:

$$T = \begin{bmatrix} 2.5653e\text{-}1 & 3.8702e\text{-}1 & 7.9427e\text{-}1 \\ 4.2417e\text{-}1 & 2.5035e\text{-}1 & 1.7708e\text{-}1 \\ 3.6682e\text{-}8 & 2.4582e\text{-}4 & 4.1504e\text{-}3 \end{bmatrix}$$

The inverse color conversion can also comprise a 3×3 matrix multiplier, with values of the inverse matrix of T. It can be seen above that T is non-singular and thus always has an inverse.

Alternative embodiments of the color converter 12 can take into account spatially distributed information, and thus compute the three color components of any image pixel based not only on its R,G,B values but also on the R,G,B values of its neighbors, as well as on color statistics of the whole image.

DECORRELATING TRANSFORM CALCULATOR 80

The decorrelating transform reduces to a minimum the cross-correlation between the various color components and/or the spatially close pixel elements of the image signal. In an example embodiment, we employ the Karhunen-Loeve Transform (KLT) to decorrelate the color components of the

image signal. Given that the color of each pixel can be represented by three elements, the KLT is a linear transform in the form of a 3×3 matrix Φ which multiplies the tricolor vector. The Transform Calculator 80 computes the values of the 3×3 transform matrix Φ.

In another example, a similar technique is used to decorrelate also spatial correlations of the colors of neighboring pixels.

The calculation comprises four steps:

(a) Computation of the mean color vector M;

(b) Computation of the autocovariance matrix C;

(c) Computation and normalization of the eigenvalues of C, which constitute the rows of Φ;

(d) Computation of the inverse Φ^{-1} (in [W92] it is shown that the inverse always exist).

(a) Computation of the mean color vector M

The ensemble of color vectors X consists of all the color triplets (e.g. V_l, V_m, V_s) over all image pixels. To simplify nomenclature, let:

$$X = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Thus, there are LxW triplets for an LxW image. The mean color vector is:

$$M = \left(\frac{1}{L \times W}\right) \sum_{\substack{all \\ image \\ pixels}} X = \begin{bmatrix} m_a \\ m_b \\ m_c \end{bmatrix}$$

This is computed by summing each color component over the whole image and diving the sum by the number of pixels. It is well known how to devise hardware to perform this computation.

(b) Computation of the autocovariance matrix C

The autocovariance matrix C is defined as follows:

$$C = E[(X-M)(X-M)^t] = E[XX^t] - MM^t$$

Substituting:

$$XX^t = \begin{bmatrix} a \\ b \\ c \end{bmatrix} [a\ b\ c] = \begin{bmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{bmatrix}$$

$$MM^t = \begin{bmatrix} m_a^2 & m_a m_b & m_a m_c \\ m_a m_b & m_b^2 & m_b m_c \\ m_a m_c & m_b m_c & m_c^2 \end{bmatrix}$$

we get:

$$C = \left(\frac{1}{L \times W}\right) \sum_{\substack{all \\ image \\ pixels}} XX^t - MM^t$$

$$= \left(\frac{1}{L \times W}\right) \sum_{\substack{all \\ image \\ pixels}} \begin{bmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{bmatrix} - \begin{bmatrix} m_a^2 & m_a m_b & m_a m_c \\ m_a m_b & m_b^2 & m_b m_c \\ m_a m_c & m_b m_c & m_c^2 \end{bmatrix}$$

$$\doteq \begin{bmatrix} d & e & f \\ g & h & i \\ j & k & l \end{bmatrix}$$

(c) Computation and normalization of the eigenvalues of C

It is well known that in order to find the eigenvalues of C we must solve for all roots λ of

$$\text{determinant}\ (C - \lambda I) = 0$$

where I is the 3×3 unit matrix. In other words, solve for all three roots of the cubic equation:

$$-\lambda^3 + (d+h+l)\lambda^2 + (ge+fj+ik-dh-dl-hl)\lambda + (dhl+eij+fgk-fhj-egl-dik) = 0$$

Since C is real and symmetric (namely $C=C^t$), all three solutions are real. Two alternative computation techniques exist for the solutions of cubic equations:

Solution 1:

Given the cubic equation in simplified form,

$$\lambda^3 + a\lambda^2 + b\lambda + c = 0$$

let

$$Q=\frac{3b-a^2}{9}$$
$$R=\frac{9ab-27c-2a^3}{54}$$
$$S=\sqrt[3]{R+\sqrt{Q^3+R^2}}$$
$$T=\sqrt[3]{R-\sqrt{Q^3+R^2}}$$

then the solutions are:

$$\lambda_1=S+T-\frac{1}{3}a$$
$$\lambda_2=-\frac{1}{2}(S+T)-\frac{1}{3}a+\frac{1}{2}\sqrt{3}(S-T)i$$
$$\lambda_3=-\frac{1}{2}(S+T)-\frac{1}{3}a-\frac{1}{2}\sqrt{3}(S-T)i$$

The implementation of this computation is straightforward, and can be devised by interconnecting adders, subtractors, multipliers, dividers, square root and cube root computing elements, and lookup tables, according to the formulae above.

Solution 2:

Given that all our roots are real and unequal, then $Q^3+R^2 < 0$. Let

$$\cos\theta = \frac{R}{\sqrt{-Q^3}}$$

The three roots can also be computed as:

$$\lambda_1 = 2\sqrt{-Q}\cos\left(\frac{1}{3}\theta\right) - \frac{1}{3}a$$
$$\lambda_2 = 2\sqrt{-Q}\cos\left(\frac{1}{3}\theta + 120°\right) - \frac{1}{3}a$$
$$\lambda_3 = 2\sqrt{-Q}\cos\left(\frac{1}{3}\theta + 240°\right) - \frac{1}{3}a$$

This computation can also be devised as an interconnection of adders, subtractors, multipliers, dividers, square root computing elements and possibly lookup tables for the cosine function.

Having found the eigenvalues λ_1, λ_2, and λ_3, we now compute the three eigenvectors as follows: For each of the roots λ_n (n=1,2,3), we compute

$$V_n = \begin{bmatrix} V_{n1} \\ V_{n2} \\ V_{n3} \end{bmatrix} = \begin{bmatrix} \frac{f(h-\lambda_n)-ei}{eg-(d-\lambda_n)(h-\lambda_n)} \\ \\ \frac{-i-gV_{n1}}{h-\lambda_n} \end{bmatrix}$$

This computation can be devised by simple interconnection of adders, subtractors, multipliers and dividers.

As the last step, the eigenvectors are normalized by dividing each element by the norm of the corresponding vector:

$$\overline{V_{ni}} = \frac{V_{ni}}{\sqrt{V_{n1}^2+V_{n2}^2+V_{n3}^2}}$$

The rows of the transform matrix Φ are these eigenvectors, ordered such that the corresponding eigenvalues λ_n are in descending order.

(d) Computation of the inverse Φ^{-1}

The computation of the inverse of the 3×3 matrix Φ is well known and can be performed by an interconnection of adders, subtractors, multipliers, and dividers.

DECORRELATING TRANSFORMER 82

One embedding of the decorrelating transformer is a 3×3 matrix multiplier, which converts the three color components of each pixel to another set of three components, decorrelated. This is especially applicable in the case that the decorrelating transform is the KLT described above.

It will be appreciated that the KLT, when applied to many types of images in the manner described above, produces three components which strongly relate to the achromatic image contents, to the red-green chromatic data, and to the blue-yellow chromatic data, respectively, in a manner reminiscent of the opponent color mechanism of the human vision system, as described for example by David H. Hubel in "Eye, Brain and Vision", Scientific American Library, 1988. Thus, the three components are also referred to as one achromatic "channel" and two chromatic ones.

Another embodiment takes into account a complete neighborhood around a pixel, and produces a converted triplet of chromatic values for the pixel.

Other embodiments perform non-linear transforms, namely, more complex transforms which cannot be achieved by simple matrix multiplications.

ENHANCEMENT PROCESSOR 40

The enhancement processor can apply any desirable enhancement to each of the three color components of each pixel. Alternative methods for enhancements exist.

One example is the chromatic enhancement as described in [W92]. It applies high-pass spatial filtering to each of the components. As a result, details are easier to observe and colors are turned more vivid. As is widely known, high-pass spatial filtering of two-dimensional signals can be achieved in a number of ways, including linear convolutions.

Another example, also explained in [W92], is chromatic enhancement in the presence of strong noise. In such a case, since empirically it was found that most noise data is confined to the achromatic channel, only the two chromatic channels are enhanced, while the achromatic channel is either low-pass filtered

or is left as is.

In a third example, each of the chromatic channels would undergo a different transform, depending not only on itself but also on the other two channels, in order to change the color of objects, so as to make the image more readily visible to people having certain degrees of "color blindness". For instance, gradients on the red-green channel would be detected and transferred to the other two channels for people with red-green color blindness.

INVERSE DECORRELATING TRANSFORMER 50 & INVERSE COLOR CONVERTER 60

Following enhancement in the transform domain, the image is transformed back to the original color coordinate. This is done by applying the inverse of two transforms (Φ^{-1} and T^{-1}).

CLAIMS

1. Color image enhancement apparatus comprising:

apparatus for receiving signals representing a color image;

image processing apparatus, employing the received signals, for image processing of the high spatial frequency chromatic components of the color image; and

apparatus for providing a color image using the output of said image processing apparatus.

2. Apparatus according to claim 1 wherein said apparatus for providing a color image also employs an achromatic component of the color image which was not processed by said image processing apparatus.

3. Apparatus according to claim 1 or claim 2 wherein said apparatus for providing a color image also employs an achromatic component of the color image which was processed by said image processing apparatus.

4. Apparatus according to any of the preceding claims and wherein said apparatus for providing a color image also employs an achromatic component of the color image.

5. Apparatus according to any of claims 1-4 wherein said apparatus for receiving comprises:

apparatus for receiving signals representing a color

image in a first color space; and

apparatus for transforming the received signals from the first color space into a color space which simulates the cones of the human visual system.

6. Apparatus according to claim 5 wherein the first color space comprises an RGB space.

7. Apparatus according to claim 5 or claim 6 wherein the cone-simulating color space comprises a (V_l, V_m, V_s) color space.

8. Apparatus according to any of claims 5 - 7 and wherein said apparatus for providing comprises apparatus for transforming the output of the image processing apparatus from the cone-simulating color space to a second color space.

9. Apparatus according to claim 8 wherein said second color space comprises an RGB space.

10. Apparatus according to any of claims 1 -9 wherein the apparatus for receiving comprises apparatus for at least partially decorrelating signals representing the color image.

11. Apparatus according to claim 10 wherein said apparatus for decorrelating comprises apparatus for performing a Karhunen-Loeve transform.

12. Apparatus according to claim 11 wherein the apparatus for providing comprises apparatus for performing an inverse of the Karhunen-Loeve transform.

13. Color image enhancement apparatus comprising:

apparatus for image processing of the high spatial frequency chromatic components of a color image.

14. Color image enhancement apparatus comprising:

apparatus for image processing of the high spatial frequency chromatic components of a color image, whereby a resulting enhanced color image is produced at least mainly from said high spatial frequency components.

15. Color image enhancement apparatus according to any of the preceding claims and wherein said high spatial frequency components comprise the top half of the spatial frequency range of the color image.

16. Color image enhancement apparatus according to any of the preceding claims 1 - 14 and wherein said high spatial frequency components comprise the top third of the spatial frequency range of the color image.

17. Color image enhancement apparatus according to any of the preceding claims 1 - 14 and wherein said high spatial frequency components comprise the top quarter of the spatial frequency range of the color image.

18. Color image enhancement apparatus according to any of the preceding claims and wherein the power spectrum of said high spatial frequency chromatic components is similar to that of a high spatial frequency achromatic component of said color image.

19. Color image enhancement apparatus for modifying a color image for perception by a color-blind individual, the apparatus comprising:

apparatus for receiving signals representing the color image;

image processing apparatus comprising apparatus, employing the received signals, for modifying the color image such that at least one color in the color image which a color-blind individual cannot differentiate is transformed to at least one color which the color-blind individual can differentiate; and

apparatus for providing a color image using the output of said image processing apparatus.

20. Apparatus according to claim 19 and also comprising apparatus for identifying at least one color which the color-blind individual does not differentiate.

21. Apparatus according to claim 19 or claim 20 wherein said apparatus for modifying comprises apparatus, employing the received signals, for modifying the color image such that a color in the color image which a dichromate does not differentiate is

transformed to at least one color which the dichromate can differentiate.

22. Apparatus according to any of claims 19 - 21 wherein said apparatus for modifying comprises apparatus, employing the received signals, for modifying the color image such that at least one color in the color image which a monochromate does not differentiate is transformed to a color which the monochromate can differentiate.

23. Apparatus according to any of claims 1 - 22 wherein said image processing apparatus comprises apparatus for enhancing chromatic differences between spatially adjacent colors.

24. Apparatus according to claim 23 wherein said apparatus for enhancing chromatic differences comprises apparatus for enhancing chromatic differences between spatially adjacent colors which are achromatically indifferentiable.

25. Apparatus according to claim 23 or claim 24 wherein said apparatus for enhancing chromatic differences comprises apparatus for enhancing chromatic differences between spatially adjacent colors which are achromatically differentiable.

26. Apparatus according to any of claims 23 - 25 and also comprising apparatus for enhancing primarily a high spatial frequency portion of an achromatic component of the color image.

27. Apparatus according to any of claims 23 - 25 and also comprising apparatus for enhancing in the presence of noise comprising apparatus for enhancing primarily a low spatial frequency portion of an achromatic component of the color image.

28. A method for color image enhancement comprising the steps of:

receiving signals representing a color image;

employing the received signals for image processing of the high spatial frequency chromatic components of the color image; and

providing a color image using the output of said apparatus for image processing.

29. A method according to claim 28 wherein said step of providing a color image also employs an achromatic component of the color image which was not processed in said image processing step.

30. A method according to claim 28 or claim 29 wherein said step of providing a color image also employs an achromatic component of the color image which was processed in said image processing step.

31. A method according to any of the preceding claims 28 - 30 and wherein said step of providing a color image also employs an achromatic component of the color image.

32. A method according to any of claims 28 - 31 wherein said step of receiving comprises the steps of:

receiving signals representing the color image in a first color space; and

transforming the received signals from the first color space into a color space which simulates the cones of the human visual system.

33. A method according to claim 32 wherein the first color space comprises an RGB space.

34. A method according to claim 32 or claim 33 wherein the cone-simulating color space comprises a (V_l, V_m, V_s) color space.

35. A method according to any of claims 32 - 34 and wherein said step of providing comprises the step of transforming the output of the employing step from the cone-simulating color space to a second color space.

36. A method according to claim 35 wherein said second color space comprises an RGB space.

37. A method according to any of claims 28 - 36 wherein the step of receiving comprises the step of at least partially decorrelating signals representing the color image.

38. A method according to claim 37 wherein said step of

decorrelating comprises the step of performing a Karhunen-Loeve transform.

39. A method according to claim 38 wherein the step of providing comprises the step of performing an inverse of the Karhunen-Loeve transform.

40. A color image enhancement method comprising the step of:

image processing of the high spatial frequency chromatic components of a color image.

41. A color image enhancement method comprising the step of:

image processing of the high spatial frequency chromatic components of a color image,

whereby a resulting enhanced color image is produced at least mainly from said high spatial frequency components.

42. A color image enhancement method according to any of the preceding claims 28 - 41 and wherein said high spatial frequency components comprise the top half of the spatial frequency range of the color image.

43. A color image enhancement method according to any of the preceding claims 28 - 41 and wherein said high spatial frequency components comprise the top third of the spatial frequency range of the color image.

44. A color image enhancement method according to any of the preceding claims 28 - 41 and wherein said high spatial frequency components comprise the top quarter of the spatial frequency range of the color image.

45. A color image enhancement method according to any of the preceding claims 28 - 44 and wherein the power spectrum of said high spatial frequency chromatic components is similar to that of a high spatial frequency achromatic component of said color image.

46. A color image enhancement method for modifying a color image for perception by a color-blind individual, the method comprising the steps of:

receiving signals representing the color image;

image processing the color image, comprising the step of employing the received signals for modifying the color image such that at least one color in the color image which a color-blind individual cannot differentiate is transformed to at least one color which a color-blind individual can differentiate; and

providing a color image using the output of said apparatus for image processing.

47. A method according to claim 46 and also comprising the step of identifying at least one color which the color-blind individual does not differentiate.

48. A method according to claim 46 or claim 47 wherein said step of employing comprises the step of employing the received signals for modifying the color image such that a color in the color image which a dichromate does not differentiate are transformed to at least one color which the dichromate can differentiate.

49. A method according to any of claims 46 - 48 wherein said step of employing comprises the step of employing the received signals for modifying the color image such that at least one color in the color image which a monochromate does not differentiate are transformed to a color which the monochromate can differentiate.

50. A method according to any of claims 28 - 49 wherein said step of image processing comprises the step of enhancing chromatic differences between spatially adjacent colors.

51. A method according to claim 50 wherein said step of enhancing chromatic differences comprises the step of enhancing chromatic differences between spatially adjacent colors which are achromatically indifferentiable.

52. A method according to claim 50 or claim 51 wherein said step of enhancing chromatic differences comprises the step of enhancing chromatic differences between spatially adjacent colors which are achromatically differentiable.

53. A method according to any of claims 50 - 52 and also comprising the step of enhancing primarily a high spatial frequency portion of an achromatic component of the color image.

54. A method according to any of claims 50 - 52 and also comprising a step of enhancing in the presence of noise comprising the step of enhancing primarily a low spatial frequency portion of an achromatic component of the color image.

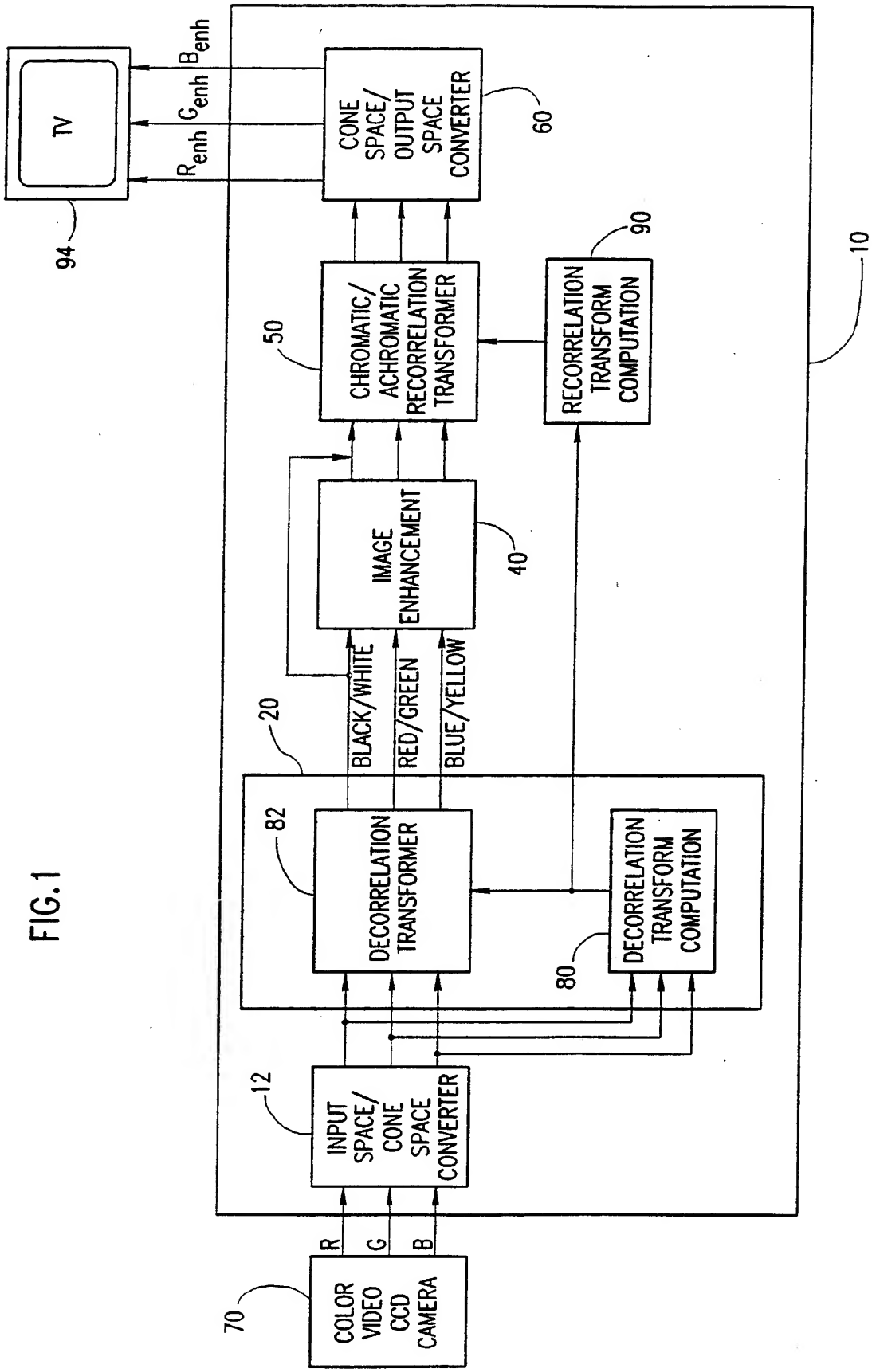

TV
94
B_{enh}
G_{enh}
R_{enh}
CONE SPACE/ OUTPUT SPACE CONVERTER
60
CHROMATIC/ ACHROMATIC RECORRELATION TRANSFORMER
50
RECORRELATION TRANSFORM COMPUTATION
90
IMAGE ENHANCEMENT
40
BLACK/WHITE
RED/GREEN
BLUE/YELLOW
20
DECORRELATION TRANSFORMER
82
DECORRELATION TRANSFORM COMPUTATION
80
INPUT SPACE/ CONE SPACE CONVERTER
12
R
G
B
COLOR VIDEO CCD CAMERA
70
10

FIG. 1

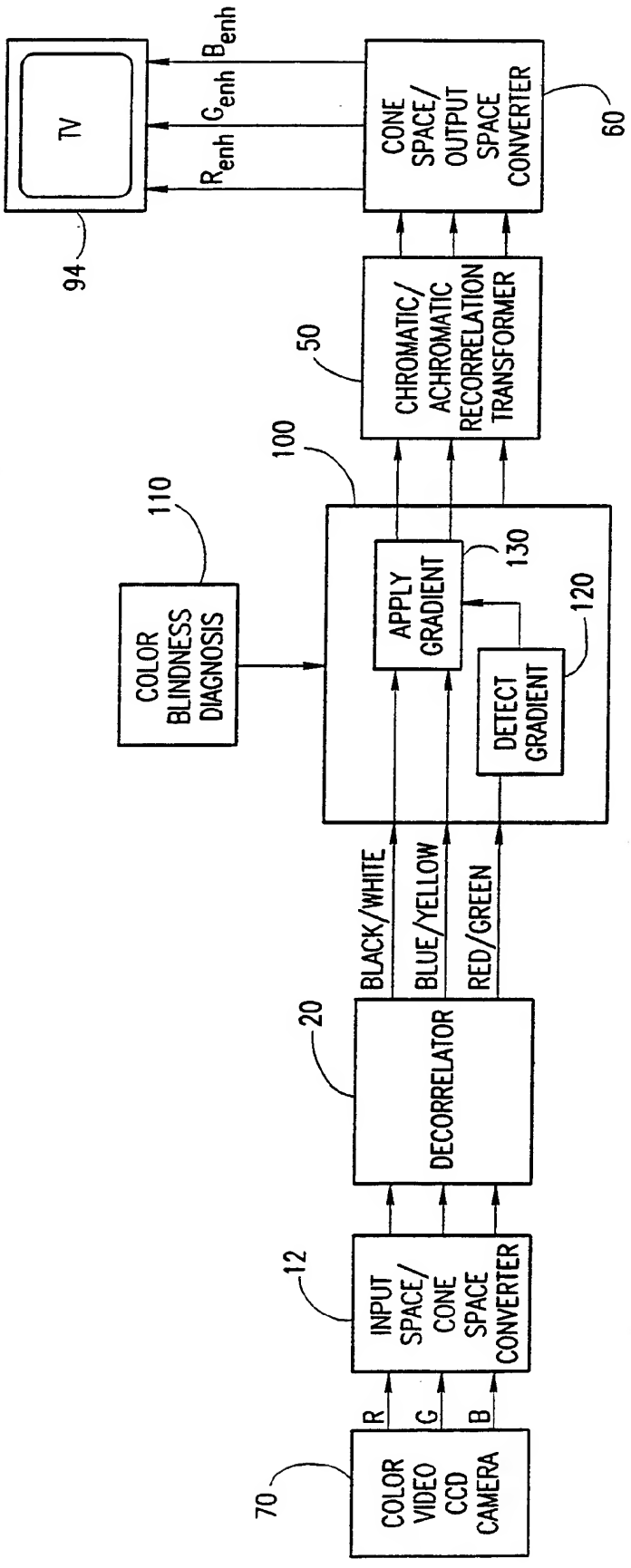

TV
94
B_{enh}
G_{enh}
R_{enh}
CONE SPACE/ OUTPUT SPACE CONVERTER
60
CHROMATIC/ ACHROMATIC RECORRELATION TRANSFORMER
50
100
110
COLOR BLINDNESS DIAGNOSIS
APPLY GRADIENT
130
DETECT GRADIENT
120
BLACK/WHITE
BLUE/YELLOW
RED/GREEN
DECORRELATOR
20
INPUT SPACE/ CONE SPACE CONVERTER
12
R
G
B
COLOR VIDEO CCD CAMERA
70

FIG.2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/11146

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :HO4N 1/46, 7/00; GO9G 1/28
US CL :358/81, 94; 345/154
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 358/81, 94, 133, 141, 538; 345/150, 154; 395/131

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS: Chromatic, Achromatic, Color Process, Power Spectrum, Karhunen-Loeve, Color Transformation, Color Space, Color Blind

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US, A, 5,079,621 (Daley et al.) 07 January 1992, col. 2, lines 5-25.	1, 13, 14, 28, 40, 41 --- 2, 3, 19, 20, 21, 29, 30, 46, 47, 48
X --- Y	US, A, 4,752,827 (Cassagne et al.) 21 June 1988, col. 2, lines 8-14.	1, 13, 14, 28, 40, 41 --- 2, 3, 19, 20, 21, 29, 30, 46, 47, 48

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 01 February 1994	Date of mailing of the international search report MAR 15 1994
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. NOT APPLICABLE	Authorized officer STEVEN J. SARAS Telephone No. (703) 305-4718

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/11146

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 4,980,760 (Hiratsuka et al.) 25 December 1990, col. 2, lines 3-17.	2, 3, 20, 21, 29, 30, 47, 48
Y	US, A, 5,125,046 (Siwoff) 23 June 1992, col. 4, lines 49-56; col. 5, line 67 to col. 6, line 29; col. 13, lines 38-53.	19, 20, 21, 46, 47, 48

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/11146

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 4-12, 15-18, 22-27, 31-39, 42-45, 49-54
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.